Elle Edit View Go Communica			18.4 X (1) 14.	VIPO FI
Bookmarks & Locatio	http://victoria.inpharmatica.co	uk/~volker/BPD3target.html		
	Target Mini Interface	inpl	famha	tica
Select Your Que	ery Sequence			
Enter PDB accession     OR	number (e.g. 1QMA): [] lfa	and chain (e.g. B): [a		
	accession (e.g. P27504) o	r GenBank proteinID (e.	g. CAB08761.1	):[]
Select Database				
Release: DEVF9 = E	BPD3□			
Apply Filters				
• Iteration Filter: PSI-B	LAST matches to be exclud	ed:		
Matches detected du	ring the first 20 forward ite	rations 🗆		
detected after 4 or more in using Genome-Threader PSI-BLAST iterations and	es delected during the first 3 it nnotation). This allows you to a SI-BLAST iterations. Matche are not effected by this option it by Genome-Threader it will SPECIES:	ocus on more remote homes s detected using PSI-BLA However, if one match is	ologous which h	ave been
☐Homo sapiens	☐Rattus norvecicus (Rat)	☐Mus musculus (Mouse)	☐Danio rerio	(Zebra fish) ▼
P=0= 100%	ALLEY BY HE TO THE	P. St. 847 1 "1 5 81563.110		\$ 'B' 60' 60' 2

# FIG. 2A

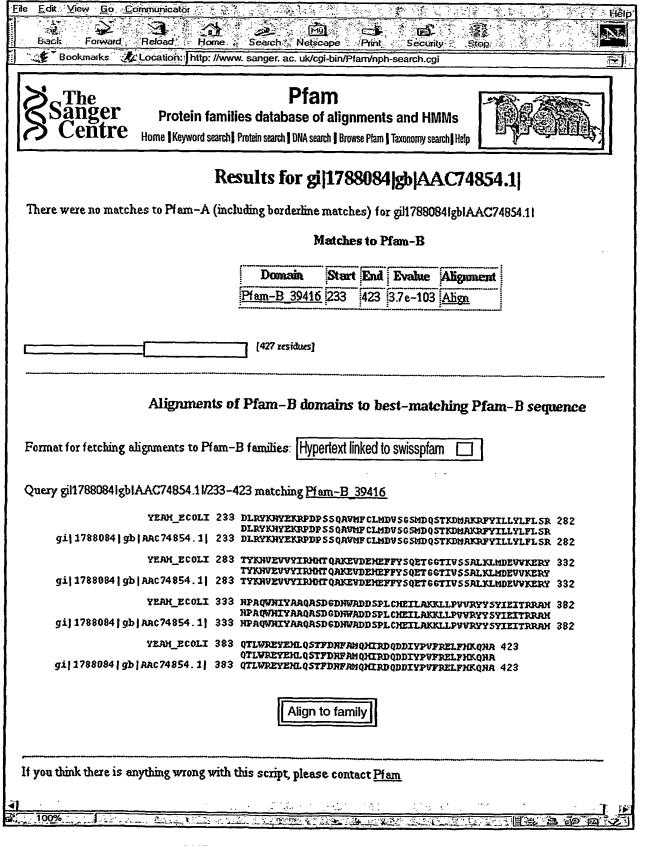
			by both, Genome Threader			AST:							
ld2lis	t EPD links	WWWlink	Title	Organism	Div.	% ID (GT,PSI)	Query rgn. (DT,PSI)	Target rgn. (GT,PSI)	Ain.	Conf. (GT)	1st Iter.	Beat Iter.	8 E-
J	AAAS9544.1 drill through Top50BlastHits	AA A 59544.1	Not given	Homo sapiens	PRI	33.7%, 33% wnmasked5W	4-183,	150-336, 150-336	449	100% unmaskedGT	(PSI)	(FSI) 2	3E
<b>&gt;</b>	Red.Seq.Vlew  AAB24821.1 drill through Top508lastHits	AAB24821.1	leukocyte integrin alpha chain	Homo sapiens	PRI	33.7%, 33% arrmaskedSW	4-183, 4-183	150-336, 150-336	449	100% unmaskedGT	1	2	3E
ــــــــــــــــــــــــــــــــــــــ	Q99715 drill through Top50 Blast Hits Red. Seq. View	Q99715	COLLAGEN ALPHA I(XII) CHAIN PRECURSOR.	Homo sapiens (Human).	PRI	28.8%, 26% arranaskedSW	4-181, 2-174	140-318, 2321-2495	440	100% unmaskedGT	1	3	2E
.ī	P20701 thill through Top50 Blast Hits Red. Seq. View	<u>P20701</u>	LEUKOCYTE ADHESION GLYCOPROTEIN LFA-1 ALPHA CHAIN PRECURSOR (LEUKOCYTE FUNCTION ASSOCIATED MOLECULE 1, ALPHA CHAIN)(CD11A)	Homo sapiens (Human).	PAI	100%, 100% ummaskedSW	1-183, 1-183	153-335, 153-335	423	100% umaskedGT	1	1	1E·
L	AAC31672.1 drill through Top50BlastHits Red.Seq.View	AAC31672.1	(INTEGRIN ALPHA-L).  leukocyte function-associated molecule-1 alpha subunit	Homo sapiens	PRI	99.5%, 99% Urimasked5W	1-163, 1-183	153-335, 153-335	423	100% unmaskedGT	1	1	16
د	CAA72402.1 drill through Top50Blas1Hits Red Seq. View	CAA72402.1	collagen type XIV	Humo sapiens	PRI	29.1%, 29% UnmaskedSW	2-180, 2-180	5-185, 5-185	422	100% unmasked GT	1	2	2E
ر	AAB38702.1 drill through Top50BlastHils Red,Seq,View		cartilage matrix protein	Homo sapiens		31.7%, 27% unmaskedSW	4-183, 2-182	275-455, 39-223	413	100% urmaskedGT	ı	2	2E
<u>۔۔۔</u>	CAB70853.) drill through Top50BlastHits Red.Seg.View		hypothetical protein	Homo sapiens	PRI	<u>28%</u> , 28% urumaskedSW	1-180, 1-183	437-620, 437-624	406	100% urmaskedGT	ı	2	1E
ے۔۔۔	CAA27972.1 drill through Top50BlastHits Red.Seq.View	CAA27972.1	Not given	Homo sapiens	PRI	20.5%, 20% unmaskedSW	3-183, 2-181	1497-1673, 1689-1873		100% urmaskedGT	2	3	1E
٦	AAB59512.1 drill through Top50BlastHils Red.Seq.View	AAB59512.1	Not given	Homo sapiens		20.5%, 20% ummaskedSW		758-934, 950-1134	405	100% unmasked GT	2	3	ΙE
د.	CAA07569.1 drift through Top50 Blast Hits	CAA07569.1	mabriin—4	Homo sapiens	PRI	28.1%, 25% unmasked5W	1-183, t-183	342-528, 31-217	403	100% urmaskedGT	1	2	2E

Helb	Z	DOL		84 Y	albulk				LTH		i i i i i	u din				
	E	The second secon	Conf.	100% unmaskedGT	99.18% unmaskedGT	99.18% unmaskedGT	98.86% unmaskedGT	98.52% unmaskedGT	98.16% unmaskedGT	97.07% unmaskedGT	95.09% unmaskedGT	93.33% unmaskedGT	92.08% unmaskedGT reverse Hit	91.34% unmaskedGT reverse Hit	91.34% unmaskedGT reverse Hit	91.34% 4 <u>unmaskedGT</u>
6 T			Aln.	123	82	8	82	78	79	79	82	8	72	74	2	7
		A Table of the Control of the Contro	Target rgn.	1-80	20-124	61 -225	318-422	1836-1950	10-137	4-131	250-335	2-90	111-251	138-277	138-277	138-277
		1	Query rgn.	109-179	7-117	5-164	7-117	5-114	5-112	5-113	5-99	63-156	3-112	3-112	3-112	3-112
			dı%	13.8% unmaskedSW 109-179	WSbakemny	11% unmaskedSW	9.6% unmaskedSW	16.4% unmaskedSW	13.2% unmaskedSW	13.8% unmaskedSW	16.7% unmaskedSW	16.2% unmaskedSW	18.9% unmaskedSW	19.6% unmaskedSW	19.6% unmaskedSW	19.6% WSP9484000
			Dj.	<b>E</b>	<b>₽</b>	<b>E</b>	PR	E	PR	PRI	BCT	PR	P.8	PBI	88	PRI .
	BPD3.pl		Organism	Homo sapiens	Escherichia coli	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens						
	on-bridge.inpharmatica.co.uk/cgi-bin/volker/getTargetBPD3.pl	only:	ТПР	PR02769	G7c protein	basic franscription factor 2, 44 KD subunit	NG37	Not given	Not given	candidate fumor suppressor protein DICE1	orf, hypothelical protein	breast cancer suppressor candidate 1	glycoprotein IIIa	Not given	platelet membrane glycoprotein illa beta subunit	Not given
	//london-bri	e Threader	WAMY link	AAF71133.1	CAB52192.1	CAA82910.1	AAD21820.1	BAA20761.1	<u>AAA36154.1</u>	AAF03046.1	AAC74854.1	AAB60942.1	AAA67537.1	AAA52589.1	AAB71380.1	AAA35927.1
Communicator	L Go to: http://lond	1) 509 hits identified by Genome Threader only:	BPD link	AAF71133.1 drill through Top50BlastHits	CAB52192.1 drill through Top508lastHits	CAA82910,1 drill through Top508lastHits	AAD21820.1 drill through Top50BlastHits	BAA20761 1 drill through Top50BlastHits	AAA36154.1 drill through Top50BlastHits	AAF03046.1 drill through Iop50BlastHits	AAC74854.1 drill through Top50BlastHits	AAB60942.1 driil through Top50BlastHits	AAA67537.1 drill through Top50BlastHits	AAA52589.1 drill through Top50BlastHits	AAB71380.1 drill through Top508lastHits	AAA35927.1 drill through
Edit View Go C	Bookmarks 🙏	hits identifie	Redundant Sequence display	Red.Sea.View	Red.Seg.Mew	Red.Seg.Mew	Red. Seg. View	Red Sea View	Red.Sea.View	Red.Sea.View	Red. Seg. View	Red. Seg. View	Red. Sea. Mew	Red.Sea.View	Red. Seg. View	Red.Sea. View
图图	***	1) 509	Add2list	7	7	-	7	7	7	٦	1		ר -	7	7	· 停

7.2	ili e	4/40	
Helb	2		
	F	THE MARIET LEAGUE AND ADDITED AND STATEMENT OF THE STREET LINES LINES AND ADDITED AND ADDI	र्वात होता है। स्थान होता होता होता होता है।
		AMIGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	AAABSHU JABSHU J
		19 19 19 19 19 19 19 19 19 19 19 19 19 1	100 mm m
***		11. 11. 11. 11. 11. 11. 11. 11. 11. 11.	A CANADA
		200 1	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 ندن		100 100 100 100 100 100 100 100 100 100	(
		다. 선생님 사람들은 생각하는 것이 있다. 그 사람들은 생각이 있다면 되었다. 그 사람들은 생각이 있다면 되었다면 되었다면 되었다면 되었다면 되었다면 되었다면 되었다면 되었	MANAGER PROPERTY OF THE PROPER
		1000.6. 12.000.  1010.0. 12.00	THE
		204.10 20	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
		120. 2001. 1921. 1	AMOUNT ANGUNA AMOUNT ANGUNA AMOUNT ANAU AMOUNT ANAU A
		12. 4. 4. 4. 4. 4. 4. 4. 4. 4. 4. 4. 4. 4.	10 - 2000 10 - 2
		1. 10001. 1. 100	1.0034(601.00) 1.0034(601.00) 1.003(601.00)
	D3.pl		전
	ま	12134, 24400011, 24400011, 245134, 24400011, 245134, 24510011, 245	1000 1000 1000 1000 1000 1000 1000 100
	tTarg	111. 2010. AAA111-44. 11102. 11102. 11102. AAA111-4. AAA111-5. AAA511-5. AAA111-5. AAA111-5. AAA511-5. AAA	71631
	er/ge	**************************************	4 - 4000 1 - 40
	/volk	19000000000000000000000000000000000000	A LESS OF THE PROPERTY OF THE
1	g-bir	2000-00-00-00-00-00-00-00-00-00-00-00-00	18
	o.uk/c	D231441_AADD51313_AAA314GL4_CO D231441_AADD51313_AAA31223_ AAA531241_AADD5131_AAA31233_ D24_AAA312542_AAA31231_AAA23123_ AAA531241_AAA31231_AAA23131_AAA23131_AAA23123_ D24_AAA312542_AAA3231_AAA23131_AAA23131_AAA23131_AAA23131_AAA23131_AAA323131_AAA323131_AAA323131_AAA323131_AAA323131_AAA323131_AAA323131_AAA323131_AAA323131_AAA323131_AAA323131_AAA323131_AAA323131_AAA33131_AA	e iteradons'  1011-14-14-15-15-15-15-15-15-15-15-15-15-15-15-15-
	-bridge.inpharmatica.co.uk/cgi-bin/volker/getTargetBPD3.pl	identified using 'positive iterations':  DI (ACOITLE INDOM, ANDRINE, CARNELLE SANSONE)  DI (ACOITLE INDOM, ANDRINE, CARNELLE CANNONE)  S. PERE CARRESTE, ARRITOLE CANNONE, CAN	ISING "negative iterations":  ADMINIST ADADINE ADADINE ADADINE ADADINE  ADMINIST ADADINE ADADINE ADADINE ADADINE  TO CONTROL ON THE ADADINE ADADINE ADDINE  TO CONTROL ON THE ADADINE ADADINE ADDINE  THE ADDINE ADADINE ADADINE ADADINE  THE ADADINE ADADINE ADADINE ADADINE  THE ADDINE ADDINE ADDINE ADDINE ADDINE  THE ADDINE ADDINE ADDINE ADDINE ADDINE  THE ADDINE ADDINE ADDINE ADDINE ADDINE  THE ADDINE ADDINE ADDINE  THE ADDINE ADDINE ADDINE  THE
	ırmat	7, 100 (2, 7)  1, 100	2 "negad 1 "10" 12" 12" 12" 12" 12" 12" 12" 12" 12" 12
	inph	A MARINE A M	18 11 18 18 18 18 18 18 18 18 18 18 18 1
1.4	ridge	100 V V V V V V V V V V V V V V V V V V	Citied Court of Court
		86 106 87 207 207 207 207 207 207 207 207 207 20	F. F. G. B.
	Go to: http: //london	705 W.	20 A 62 A 63 A 64
jo;	thitp:		The second secon
nunica	Go (c	LAST LEGANGE CONTROLL C	AST 1  CARSO  OME J  OM
Edit View Go Communicator	*	out of these 645 PSI-BLAST matches were identify the state of these 645 PSI-BLAST matches were identify the state of the s	Genome Threader Matches:  Genome Threader Matches:  Genome Threader Matches:  Matches found by Genome Threading:  Genome Threader Matches:  Genome Threader Matches:  Matches found by Genome Threading:  Genome Threader Matches:  Anders:  Ande
යි	arks		Coase ound!
Vie₩	Bookn	these	these chests and the second the s
	Bookmarks &	630 out of these 645 PSI-BLAST matches were 630 out of these 645 PSI-BLAST matches were children in the control of the control	15 out of these 645 PSI-BLAST matches were identified using 'negative iterations':  Paritity Loadities Loasies Cariffel Antificial Antificial Cariffel Antificial Ant
읦		ĬŠ	15 out o A2) G4

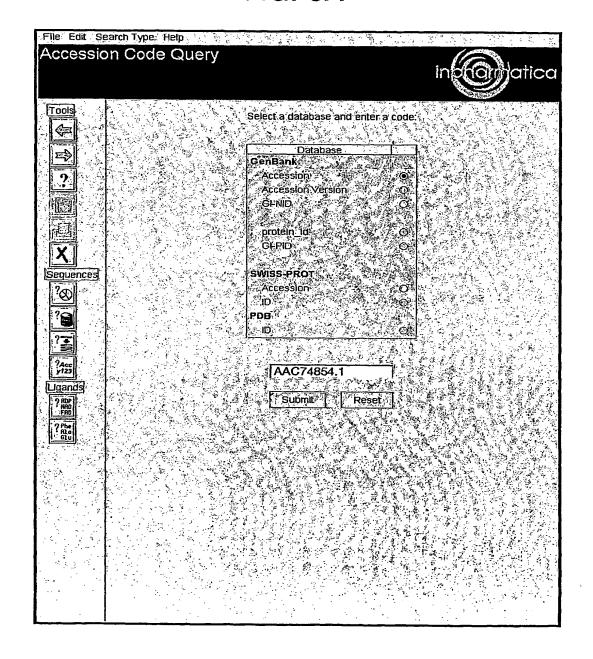
FIGUR Jatica	annotation) Code	BAA15585.1 ← P76235	AAC/4854.1 S	0		
Redundant Search Type Help Redundant Sequence Display Contains: 2 Sub-sequences. 0 PROSITE hits. 0 PRINTS hits.	Tools   Representative code: BAA15585.1 Length: 427 Organism name: Escherichia coli [additional annotation]   (全)   Aligned sequences   Code   Code		Primary database information	Sequences   BAA15585.1   Tax ID:562   Organism name:Escherichia coll   Coll	SWISS-PROT  Code Details  P76235 Tax ID:562  EC Number:Not available  Description:HYPOTHETICAL 49.4 KDA PROTEIN IN GAI	Start     個 😂 🔄   🕲 BPD4-FE1.4   影Ingharmati :: 《SFegan, Richa   与Inbox - Outlo   版 Target Minh   图Microsoft Po   図

FIG



<u>File</u> <u>E</u> dit	<u>Yiew Go Communicator</u> Ho	èlp
<b>1</b>	Bookmarks 🚜 Location: http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?cmd=Retrieve&db=🔽	JaT,
LOCUS DEPINITION ACCESSION PID	AAC74854 427 aa BCT 01-DEC-2000 orf, hypothetical protein [Escherichia coli K12]. AAC74854 g1788084	
VERSION DBSOURCE KEYWORDS	AAC74854.1 6T: 1788084 locus AE000273 accession <u>AE000273.1</u>	
SOURCE ORGANISM	Escherichia coli K12. <u>Escherichia coli K12</u> Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  Bscherichia.	
REFERENCE AUTHORS	1 (residues 1 to 427) Blattner,F.R., Plunkett,G. III, Blooh,C.A., Perna,H.T., Burland,V., Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F., Gregor,J., Davis,H.W., Kirkpatrick,H.R., Goeden,M.R., Rose,D.J., Mau,B. and Shao,Y.	
TITLE JOVRNAL MEDLIKE	The complete genome sequence of Escherichia coli K-12 Science 277 (5331), 1453-1474 (1997) 97426617	
PUBMED REPERENCE AUTHORS	9278503 2 (residues 1 to 427) Blattner,F.R.	
TITLE JOURNAL	Direct Submission Submitted (16-JRN-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA. Email: ecoli@genetics.wisc.edu Phone: 508-262-2534 Fax: 608-263-7459	
REFERENCE AUTHORS TITLE	3 (residues 1 to 427) Blattner,F.R. Direct Submission	
J OURHAL	Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA. Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459	
REFERENCE AUTHORS	4 (residues 1 to 427) Plunkett, G. III.	
TITLE JOVENAL	Direct Submission Submitted (13-0CT-1998) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA	
C OMMENT	This sequence was determined by the E. coli Genome Project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by HIH grants HG00301 and HG01428 (from the Human Genome Project and HCHGR). The entire sequence was independently determined from E. coli K12 strain MG1655. Predicted open reading frames were determined using Genemark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GR, 30332 [e-mail: mark@amber.gatech.edu]. Open reading frames that have been correlated with genetic loci are being annotated with co Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by	
	Mary Berlyn. A public version of the database is accessible (http://cgsc.biology.yale.edu). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli	
	Genome Project's World Wide Web site (http://www.genetics.wisc.edu). *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes: Aumotation updates: updated gene identifications	A COUNTY
	and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene	
PEATURES	names.  Method: conceptual translation.  Location/Qualifiers	31
source		
Protei	/db_xref="taxon:83333" n 1427	
CDS	/function="orf; Unknown" /product="orf, hypothetical protein" 1. 427 /gene="yeak"	9
	/coded_by="1788078:63857668" /trans1_table= <u>11</u> /note="0427; This 427 aa ORF is 28 pct identical (43 gaps)	
DRIGIN	to 327 residues of an approx. 312 aa protein YZDC_BACSU SW: P45742"	
121 kg	twfidrrln gknksmvnrq rflrrykaqi kqsiseaink rsvtdvdsge svsiptedis pmfhqgrgg lrhrvhpgnd hfvqndrier pqgggggsgs gqgqasqdge gqdefvfqis deyldllife dlalpnlkqn gqrqflteykt hragytangv panisvvrsl qnslarrtam agkrrelha leenlaiisn sepaqlleee rlrkeiaelr akiervpfid tfdlryknye	•
100%		

FIG. 6A



o o	[ - 1 kg			144	4	. <u></u>	 ⟨``)	س داد د د	3 %		7	ida Maria	<u> </u>			35						- 5		7.		2		T	احزاا	ìsh
Sin Jatic				Confidence	Marginal(95%)	Marginal(95%)	Marginal(95%)	Marginal(95%)	Marginal 35%)	[ DW(94%)	Law(94%)	Low(94%)	Low(94%)	Low(94%)	Low(94%)	Law(93%)	Low(93%)	Low(82%)	Low(89%)	Low(80%)	Low(72%)	Low(65%)	Law(53%)	Low(53%)	Low(21%)	Low(21%)	Law(21%)	Low(20%)		Sales of the sales of the sales
) Journal				Method	Local	Local	Local	Local	Local	le l	Local	Local	Local	Local	Local	Local	Local	Local	Loca	Lucal	Local	Local	Local	Local	Local	Local	Local	Local		A Part of the Part
				Aln. scare	78	78	78	8/2	9/	76	76	76	76	76	-70	72	78	75	75	7.6 3.88 3.	72	. 69	63	63	23	28	88	55		Table of the second
				Target rgn.	5 - 89	5 - 99	5 - 99	S .	200	4 - 128	4-128	4~128	4-128	4-128	921 - 1	5 - 101	8 - 89	4-128	4-128	5 - 83 5 - 87	1-58	18 - 68	2 - 88	5 - 99	1 ~ 105	166 226	166 - 226	2 - 86		A COLUMN TO STATE OF THE STATE
	Jument	sults:		Query rgn, I-T	250 - 335	250 - 385	<b></b> }-	250 - 335	250 - 230	250 - 365	<del>   </del>	250 - 365	250 - 365	250 - 365			250 - 931		-	250 = 333		250 - 301	245 - 331	245 - 331	242 - 335	246 - 305   1	-	250 - 331		AND STATE OF THE S
	Vièw alignment	ader re	lect all	9   Q %	-			10.7	_		+		15.1 2	-	15.1		15,3			18.07	-	21.2 2	15.2 2	15.2	13.1	13.1	13,	10,5		A STANSON
		enome Thre	Deselect all	Organism	H.sapiens	Hisapiens	H.sapiens	H.Sapiens.	Usaplens	Hsaniens	Hsaplens	H.saplens	H.saplens	H.saplens	Hsapiens	H.saplens	Hsapiens	H.sapiens	Hsaplens	Hamiens	Hsaplens	H. Saplens	H.saplens	H,sapiens	Hsaplens	Scerevisiae	Scerevisiae	<b>Anorvegicus</b>		A CANADA
	Total selected: 0.	Inpharmatica Genome Threader results:	Select all		30UND MN++	JND MAGNESI	UND MAGNESI	UND MAGNESS	COLLA I-DONIALIA WITH BOOKED MAGNESI	TAL FREE	JM COMPLEX	JM COMPLEX	MAC-1 I DOMAIN MAGNESIUM COMPLEX	TAL FREE	HUMAN VON WILLEBRAND FACTOR A3 DO.,	HUMAN VON WILLEBRAND FACTOR AS DO.,	NMR SOLUTION STRUCTURE OF THE INSER.,	-DOMAIN FROM INTEGRIN CR3, MG2+ BO	-DOMAIN FROM INTEGRIN CR3, MINZ+ BO	CD11A I-DOMAIN WITHOUT BOOM CATE	SEVETA) STRUCTURE OF THE VON WILLE	A1 DOMAIN OF YON WILLEBRAND FACTOR	A3 DOMAIN OF VON WILLEBRAND FACTOR	A BOMAIN OF YON WILLEBRAND FACTOR	DOMAIN FROM INTEGRIN ALPHAZ-BETAI	SRYSTAL STRUCTURE OF THE 20S PROTE.	CRYSTAL STRUCTURE OF THE 20S PROTE	RYSTAL STRUCTURE OF RAT A1B1 INTEG.		
	01		Total hits: 36. Selected: 0.	Title	CD11A I-DOMAIN WITH BOUND MN++	CD11A I-DOMAIN WITH BOUND MAGNESI	CD114 I-DOMAIN WITH BOUND MAGNES	COLLA I-DOMAIN WITH BOUND MAGNES	MAIN WITH BU	MAC-1   DOMAIN METAL FREE	DOMAIN CADMIUM COMPLEX	MAC-1   DOMAIN CADMIUM COMPLEX	MAIN MAGNES	MAC-1 I DOMAIN METAL FREE	WILLEBRAND	WILLEBRAND	IN STRUCTURE	ROM INTEGRIN	TOM INTEGRIN	COLIA - DOMAIN WITHOUT BOUND CATE	ALICTURE OF T	OF VON WILLE	OF VON WILLE	JE VON WILLE	OM INTEGRIN	RUCTURE OF T	RUCTURE OF T	SUCTURE OF BA		סהם
play thetical protein					-				4	MAC-1 DO	MAC-11D	MAC-11D	<u> </u>	_		-	NIMR SOLUTIO	I-DOMAIN FI	II-DOMAIN FI	C011A 1-00	CRVSTAI ST	A1 DOMAIN		-	<u> </u>	Γ,	1_			S.PROT NODR
e Disp orf, hypoth			l of 2.	Code	1LFA:A	1200:A	1200:B	1707.	V.1071		1BHQ:2	18HQ:1	18月02	11DN:2	1ATZ:B	1ATZ:A	1DGQ:A	1100	1.JLM	NO71	10AK-4	1AU@	1A09:A	1AO3.B	1A0X:B	1RYP:D	1RYP.R	1CK4B		DGanBank DSWISS
earch, type, Help Sequence Disk retails: AAC74854.1: orf, hypotl			Page 1	Detall.	A		4	1	1		1	۵	4	۵	٩	Δ	٩	Δ	۵		1	6	Γ	۵	۸	^	4	۵		Rank
dit Search Ty 1ed Sequ Query details: AAC748			R	Cluster				1	ו	+	1			5	ը		٥	L	7)	4	ונ	4	ը	j	٥	L	1	ם	Į-	765
File Edit Search Type Help Aligned Sequence Query details: AAC74854.1: orf	Tools	1		\$ 100 miles			"		N	><		Saguanhaci							?Acc	723	Ligands	280P	2	7 Phe	900			ではある。		

... G

10/40

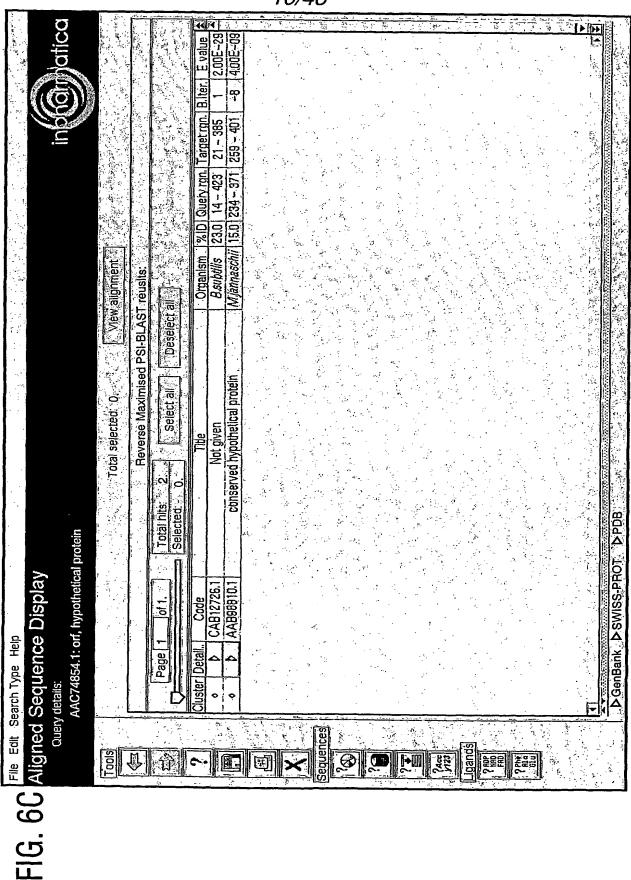
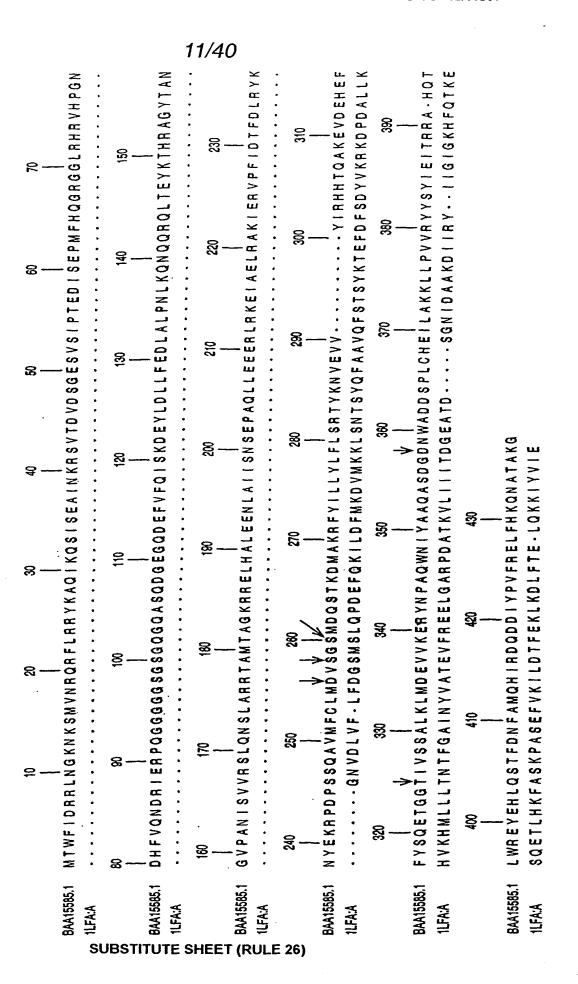
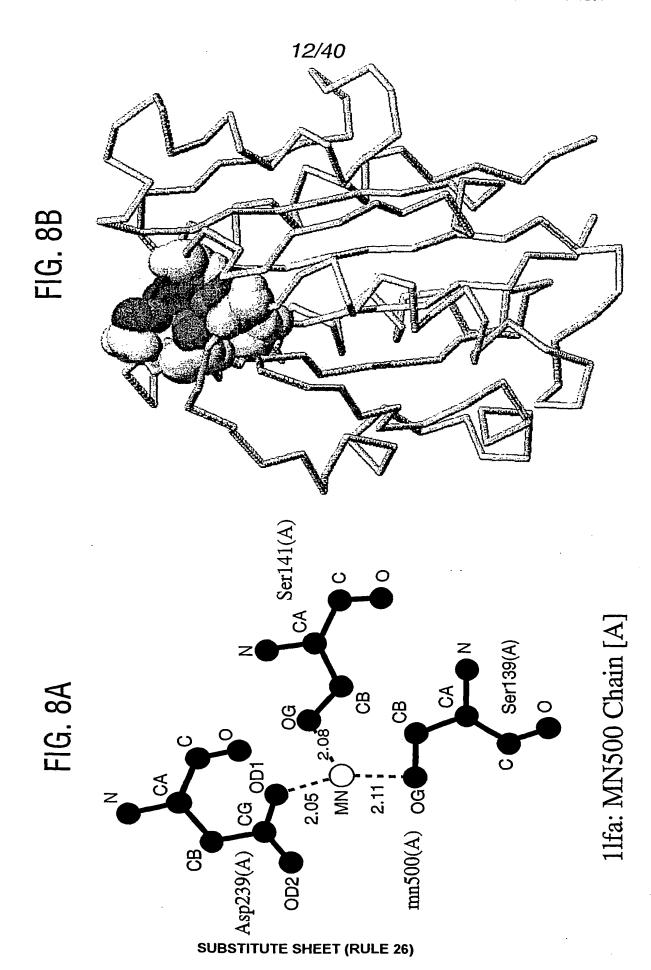


FIG. 7 AlEye output (January 4, 2002 3:07 PM)





	Communicator		* Charles and the second	Help
Bookmarks	Location: http://victoria	a. inpharmatica.co.uk/~	volker/BPD3target.html	
		Target Mining Interface	inpharmatica	
Select Yo	our Query Seque	ence		
• Enter PDB	accession number (e.g.	1 QMA): 1AOX] and chai	n (e.g. B): A]	
• Enter one	Swiss-Prot accession (e	.g. P27504) or GenBank	proteinID (e.g. CAB08761.1):	
Select Da	atabase			
Release:	DEVF9 = BPD3			
Apply Fil	ters			
• Iteration F	iter: PSI-BLAST matche	s to be excluded:		
	None	C		
report (using the been detected iterations or us	e first_PB_iter annotation after 4 or more PSI-BLAS ing Genome-Threader ar PSI-BLAST iterations and	). This allows you to foo ST iterations. Matches o e not effected by this o	these matches will be excluded from to tus on more remote homologous which letected using PSI-BLAST with negative bition. However, if one match is found do it will be excluded.	have 🎘 /e

	N	Ē	विस			(1781) '			K.L.						augi		EL
					Best E-value (PSI)	2E-73	11-11	46-58	4E-68	2E-42	1E-60	16-60	SE-6	3E-46	36-50	SE-63	18-47
					Best Her. (PSI)	2	-	2	~	2	~	~	6	~~	~	73	2   10
					lst (PSI)	- H	-	-	-	-		-	2 1	- H	-		-
					Conf. (GT)	100% unmaskedGI	100% unmaskedGI	100% unmaskedGI	100% unmaskedGT	100% urmaskedGT	1007. UndaskedGT	100% vomaskedGT	100x unmaskedQI	100% unmaskedGT	100% unmaskedGT	100% wmwskedGT	100x Immasked
					Alb. Score (GT)	487	471	463	469	637	£\$	£\$.	52	<b>4</b> 24	422	619	ŧ [
					Target rgn (GT,FSI)	139-337, 139-337	169-367, 169-367	159-355,	159-355,	133-318, 2323-2495	144-338, 144-338	144-338, 144-338	293-472, 1581-1626	271-455, 37-224	133-318, 133-318	162-359, 162-359	335-529, 28-223
					Query rgn. (GT,PSI)	2-200, 2-200	2-200, 2-200	2-200, 2-200	2-200, 2-200	2-166, 7-189	2-200, 2-200	2-200, 2-200	5-195, 156-201	3-198, 3-198	2-185, 2-195	2-189, 2-139	1-189, 2-200
					%ID (GT,FSI)	51.82, 51% vrmasked6W	99.5% 100% umasked5W	45.2%, 45% unmesked 6W	45.2%, 45% unmaskedSW	28.1%, 31% ummaskedSW	27.1%, 27% unmaskedSW	27.1%, 27% umaskedSW	26.9%, 26% unmarkedSW	26.8% 26% unmasked5W	22.6%, 30%, wnmaskedSH	46.5% 46% wmaskedgW	26%, 28% gramssked5W
					DIV.	P.R.	E	E	E	E	P.B.	E	똢	E	æ.	Ä.	₹.
	Sle	3.pl#Mine2			Grgswism	Homo sapiens (Human).	Homo saplens (Human).	Homo saplens	Homo saplens	Homo sapiens (Human).	Homo sapiens	Homo saplens	Homo saplens	Homo sapiens (Human).	Homo sapiens	Homo suplens	Homo saplens
	子 之 [2] (1) (1) (2) (2) (3) (4) (4) (5) (5) (7) (7) (7) (7) (7) (7) (7) (7) (7) (7	: //london-bridge.inpharmatica.co.uk/cgf-binvolker/gatTargetBPD3.pl#Mine2	nome Threader and PSI-BLAST:	BLAST raives are there in many on!	AIAT.	INTEGRIN ALPHA-1 (LAMININ AND COLLAGEN RECEPTOR) (VLA-1) (CD49A).	PLATELET MEMBRANE GLYCOPROTEIN IA PRECURSOR (GFIA) (COLLAGEN RECEPTOR) (INTEGRIN ALPHA-2) (VLA-2 ALPHA CHAIN) (CO49B).	Friegrin Apha-11 suburit precuraor	Integrin alpha II subunit precursor	COLLAGEN ALPHA 1(XII) CHAIN PRECURSOR.	Nat given	Not given	d.R38D15.1 (colagen, type XII, apha 1)	CARTILAGE MATRIX PROTEIN PRECURSOR (MATRILIN-1).	negalos IIX aqti	integrn subuni alpha i 0 precusor	natrin-4
ig G	Release The Part of the Part o	🎉 Location   http://london	by both, Ge	it output: PSI	WWW Ent.	201923 	P17301	AAF01258.1	AAD51919.2	299715	AAA5240.L	AAA59546.1.	CABTIZZA	क्राध्य	AAC01506.1	AAG31832.1	CAA07569.1
View 120	Forward	Bookmarks & Loc	2) 82 additional hits identified by both, Genome T	Combined Grama Threader and PSI - Blact octor; PSI - BLAST vel	Ե <b>Р</b> Ծ №1:;	PSELSE dra lincoch IopSQBisatHits Red. Seq. New	P17301 dril through Ion508iss1His Red. Sea Mew	AAF01259.1 drill through 10050BlastH1s fled.8cg.Mew	AAOS1919.2 drii throwen IooSOBiestHio Bed.Seq.Mew	G99715 dra thrown Iop50BisstHis Red Seq.Mew	AAASS919.1. dril lhrouen IonSoBlastHis Red. Seq. Mew	AAAS9544.1 drii Ibrouch Iop50Bissitiis Ard.Sen.View	CA871222.1 drii Inteuch Iop50BlastHits Bed.Seg.Vew	P21941 drift through Top508half41s Red.Seq.View	AACOISO6.1 fr# Inrovah Ioo50BlastHis. Red.Seg.Mew.	AAC31852.1 drii Urosch Ion50Bissiffils Red Sen Mew	CA407569.1 dra Inroyan Ion508041Hits
	98ck	**	2) 82 add	Combined C	!<:ld2  s1	٦	ר	٦	ה	ם	ב	ה	7	٦	7	3	7 %

			1	5/40			
	2E-8	9 <del>-</del> -9	1E-4	5E-7	3E-24	6E-22	က ည
	w.	2	ហ	ന	ι'n	<b>r</b> -	ထု
Φ	m	2	ហ	ന	4	යි	7
GT Confidence	unmaskedGT	100% unmaskedGT	100% unmaskedGT	100% unmaskedGT	100% unmaskedGT	100% unmaskedGT	98.52% unmaskedGT reverse Hit
ပ	142	121	117	114	107	102	Ξ
	36-141	3-51, 37-93	312-421, 312-421	7-92,7-92	267-384, 265-424	214-398, 208-380	145-187, 145-187
	80-195	2-46, 139-196	8-117, 8-117	108-197, 108-197	9-134, 7-186	8-192, 2-181	157–199, 157–199
10B	unmaskedSW	34.7%, 21% unmaskedSW	15%, 15% unmasked5W	21.1%, 21% unmaskedSW	13.2%, 11% unmaskedSW	12.9%, 14% unmaskedSW	14% 13% 157-199, unmaskedSW 157-199
•	<u> </u>	<u> </u>	<u> </u>	<u>E</u>	BCT	8CT	<b>E</b>
FIG.	sapiens	Homo	Homo sapiens	Homo sapiens	Escherichia coli	Escherichia coli	Homo sapiens
	Not given	undulin 2	calcium—activated chloride channel—2	collagen VI-alpha-1 chain	orf, hypothetical protein	tellurite resistance	platelet membrane glycoprotein Ilb
	<u>BAA91707.1</u>	<del>V</del>	AAD40367.1	CAA67559.1	<u>AAC76768.1</u>	CAB43000.1	<u>AAA60114.1</u>
	Top50Blast Hits BAA91707.1 Red. Seq. View.	AAA36795.1 driil through Top50BlastHits Red. Seg. View	AAD40367.1 drill through Top50BlastHits Red.Seg.View	CAA67559.1 drill through Top508lastHits CAA67559.1 Red.Seg.View	AAC76768.1 drill through Top50BlastHits Red.Seg.Mew	CAB43000.1 drill through Top50BlastHits Red. Seg. Mew	AAA60114.1 drill through Top50BlastHits Red. Seg. View
	٦	٦	7	٦	<b>†</b>	٦	7

**SUBSTITUTE SHEET (RULE 26)** 

# FIG. 10C

608 out of these 632 PSI-BLAST matches were identified using 'positive iterations':

| 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 1922 | 16/40

24 out of these 632 PSI-BLAST matches were identified using 'negative iterations':

Babilid (<u>Caridose Carbsing Carbidol Cardidie Cardidie Cardidie Cardidie Cardidie Cardise Colste Cardisel Cardisel Cardige Cardisel Cardis</u>

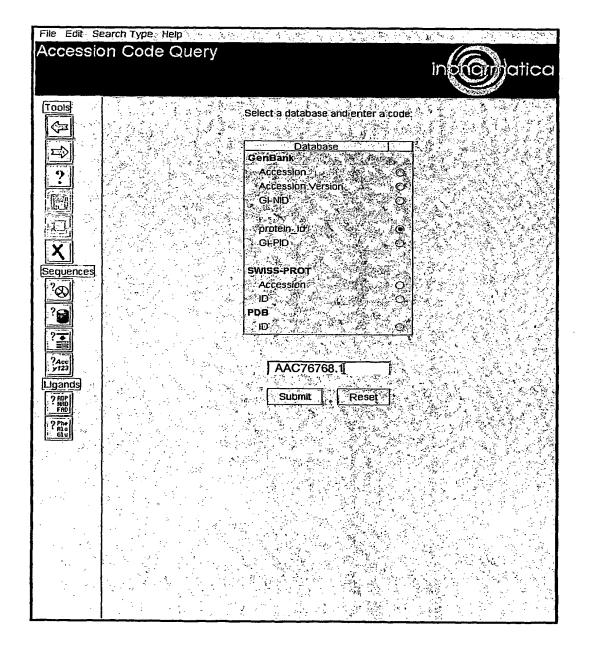
File Eff. Search Type Help
----------------------------

7

le Edit View Go Communicator He  Bookmaiks Location: http://www.sanger.ac.uk/cgi-bin/Pfarn/nph-search.cgi
Stanger Protein families database of alignments and HMMs Centre Home   Keyword search   Protein search   Browse Plam   Taxonomy search   Help
Results for gi 2367274 gb AAC76768.1
There were no matches to Pfam-A (including borderline matches) for gil23672741gblAAC76768.11
Metches to Pfean-B
Domain Start End Evalue Alignment Ptam-B 15204 204 408 2.4e-108 Align
[427 residues]
Alignments of Pfam-B domains to best-matching Pfam-B sequence
Format for fetching alignments to Pfam-B families: Hyperlext linked to swisspfam
Query gil2367274lgblAAC76768.11/204-408 matching Plam-B 15204
YIEM_ECOLI 204 DILRILPPELATLGITELEYEFYRRLVEKQLLTYRLHGESVREKVIERPV 253 DILRILPPELATLGITELEYEFYRRLVEKQLLTYRLHGESVREKVIERPV gi[2367274 gb]arc76768.1  204 DILRILPPELATLGITELEYEFYRRLVEKQLLTYRLHGESVREKVIERPV 253
YIEM_ECOLI 254 VHKDYDEQPREPFIVCVDTSCSMGGFREQCRKAFCLALMRIALAEHRRCY 303 VHKDYDEQPREPFIVCVDTSGSMGGFREQCAKAFCLALMRIALAEHRRCY gi[2367274]gb]aac76768.1] 254 VHKDYDEQPREPFIVCVDTSGSMGGFREQCAKAFCLALMRIALAEHRRCY 303
YICH_ECOLI 304 IMLFSTEIVRYELSEPQEIEQAIRFLSQQFRGETDLASCFRAIMERLQSR 353 IMLFSTEIVRYELSEPQEIEQAIRFLSQQFRGETDLASCFRAIMERLQSR gi[2367274]gb]arc76768.1 304 IMLFSTEIVRYELSEPQEIEQAIRFLSQQFRGETDLASCFRAIMERLQSR 304 IMLFSTEIVRYELSEPQEIEQAIRFLSQQFRGETDLASCFRAIMERLQSR 353
YIEM_ECOLI 354 EWFDADAVVISDFIAQRLPDDVTSKVKELQRVHQHRPHAVAMSANGKPGI 403 EWFDADAVVISDFIAQRLPDDVTSKVKELQRVHQHRPHAVAMSANGKPGI gi 2367274 gb aac76768.1  354 EWFDADAVVISDFIAQRLPDDVTSKVKELQRVHQHRPHAVAMSANGKPGI 403
YIEM_ECOLI 404 MRIPD 408 MRIPD
gi 2367274 gb Arc76768.1  404 HRIFD 408
Align to family
If you think there is anything wrong with this script, please contact <u>Pfam</u>

BOOKMERKS & LOCATION	File Edit	<u>View Go Communicator</u>
PETINITION of Dypothetical protein [Escherichia coli X12]  PERSION ARCESSES	<b>1</b> Bc	pokmarks Location: http://www.ncbi.nlm.nih.gov.80/entrez/query.fcgi?cmd=Retrieve&db=protei
ACCUSATION AND ACCUSED TO THE ACCUSATION AND ACCUSATION ASSOCIATED ASSOCIATION AND ACCUSATION ASSOCIATION ASSOCIAT		ARC76768 427 aa BCT 01-DEC-2000
PID AGGS782: CT.2367274  PREVIOUS AGGS782: CT.2367274  PREVIOUS AGGS782: CT.2367274  PREVIOUS SENDERICH COL NT.  Escherichia col Nt.  E		orr, hypothetical protein [Escherichia coli K12]
VERSION SACCESSES. 1 et. 2257274  DOUBLES  DEFERENCE  REPREMENCE  1 (residues 1 to 427)  AUTHORS  Blattner, P.R., Plumkett, E. III, Bloch, C.A., Perna, R.T., Burland, V., Reigner, D. Avie, D. L. Limpathich, K.A., Gooden, G.A., Mose, D.J., Haw, P. and Shao, V.  THILE JOHN BLATT COMMENT OF THE COMMENT OF THE COMPLETE STATES OF T		
DRSOUNCE lower ACOUSTS accession RECOUSTS!  PORRANTS  SECHETICAL SECRETICAL S		
Schorischis coli X12  BRANKISS  BRAN	DBSOURCE	
ORGANISM  Backerick; Protebosteria; gamma subdivision; Enterobacteriacea;  Backeria; Protebosteria; gamma subdivision; Enterobacteriacea;  Backeria; Protebosteria; gamma subdivision; Enterobacteriacea;  Batter, P. R., Piunkett, E. III, Bloch, C.A., Perma, R.T., Burland, V.,  Blattner, P. R., Piunkett, E. III, Bloch, C.A., Perma, R.T., Burland, V.,  Blattner, P. R., Piunkett, E. III, Bloch, C.A., Perma, R.T., Burland, V.,  Blattner, P. R., Piunkett, E. III, Bloch, C.A., Perma, R.T., Burland, V.,  Blattner, P. R., Piunkett, E. III, Bloch, C.A., Perma, R.T., Burland, V.,  Baugh, B. and Shao, Y.,  Burland, P. Barland, P. Barland, P. Burland, V.,  Burland, P. Barland, P. Bar		•
REFERENCE  AUTHORS  Bilther, P.R., Pinnkett, E. III, Bloch, C. R., Perna R.T., Burland, V., Sergor, S., Dawis, R.V., Kirpatrick, R.R., Sceden, B.A., Rose, D.J., Blither, P.R., Pinnkett, E. III, Bloch, C. R., Perna R.T., Burland, V., Scegor, S., Dawis, R.V., Kirpatrick, R.R., Sceden, B.A., Rose, D.J., Blither, P.R., Soulies, S.V., Kirpatrick, R.R., Sceden, B.A., Rose, D.J., Blither, P.R., Soulies, R.V., Kirpatrick, R.R., Sceden, B.A., Rose, D.J., Brown and C. S., School, C. R., School, R. R., Scho		
REFERENCE AUTHORS  AUTHORS Blattner, P.R., Pikmkett, G. III, Bloch, C.R., Perma, R.T., Burlind, V., Riley, B., Collado-Vides, J., clarner, J.D., Rede, C.K., Hayhew, G.F., Obacis, P.V., Kirhystrick, M.A., Coeden, B.A., Rose, D.J., DAN, P.A., Collado-Vides, J.,	ONGMITSE	Syneticals coll AIC
AUTHORS  Blattner, F. R., Plumkett, E. III, Bloch, C. R., Perma, R. T., Burland, V., Blaup, W. Collidon-Guer, J. Blarmer, J. D., Rode, C. K. Haphew, E. F., Blaup, F., and Shao, V. Ritpatrick, M. R., Goeden, B. R., Rose, D. J., Haup, P., and Shao, V. Ritpatrick, M. R., Goeden, B. R., Rose, D. J., The complete genome sequence of Escherichia coli K-12 science 277 (5331), 1453-1474 (1997)  RIPERENGE 2 (residues 1 to 427)  AVTHORS Blattner, P. R., 100  RIPERENGE 2 (residues 1 to 427)  AVTHORS Blattner, P. R., 100  RIPERENGE 3 (residues 1 to 427)  AUTHORS Blattner, F. R., 100  RIPERENGE 3 (residues 1 to 427)  AUTHORS Blattner, F. R., 100  TOURNAL Submitted (16-7MN-1937) cmp Plumkett III, Laboratory of Senetics, University of Wisconsin, 445 Henry Hall, Hadison, WI 53706, USA.  REFERENCE 4 (residues 1 to 427)  AUTHORS Blattner, F. R., 100  TOURNAL Submitted (02-2mp-1939) cmp Plumkett III, Laboratory of Senetics, University of Wisconsin, 445 Henry Hall, Hadison, WI 53706, USA.  REPREDICE 4 (residues 1 to 427)  AUTHORS Plumket, C. III.  TOURNAL Submitted (02-2mp-1939) Laboratory of Senetics, University of Wisconsin, 445 Henry Hall, Hadison, WI 53706, USA.  REPREDICE 4 (residues 1 to 427)  AUTHORS Plumket, C. III.  TOURNAL Submitted (02-2mp Hall, Hadison, WI 53706, USA.  REPREDICE 5 (residues 1 to 427)  AUTHORS Plumket, C. III.  TOURNAL Submitted (02-2mp Hall, Hadison, WI 53706, USA.  REPREDICE 6 (residues 1 to 427)  AUTHORS Plumket, C. III.  TOURNAL Submitted (02-2mp Hall, Hadison, WI 53706, USA.  REPREDICE 7 (residues 1 to 427)  AUTHORS Plumket, C. III.  TOURNAL Submitted (02-2mp Hall, Hadison, WI 53706, USA.  REPREDICE 8 (residues 1 to 427)  AUTHORS Plumket, C. III.  TOURNAL Submitted (02-2mp Hall, Hadison, WI 53706, USA.  REPREDICE 9 (residues 1 to 427)  AUTHORS Plumket 1 to 4270  AUTHORS Plumket 1 to 4270		Escherichia.
eregor, J. Davis, R. G., Kirkpatrick, H. M., Gooden, R. Rose, D. J.,  TITLE  JOURNAL,  MEDILINE  PUMMED  Direct Submission  Submitted (16-7M-1997) cmy Plunkett III, Laboratory of Senetics,  Wiley (25-745) cm, 445 Menry Hall, Hadison, WI 53706, USA.  BITTLE Submission  JOURNAL  JOUR		1 (residues 1 to 427)
eregor, J. Davis, R. G., Kirkpatrick, H. M., Gooden, R. Rose, D. J.,  TITLE  JOURNAL,  MEDILINE  PUMMED  Direct Submission  Submitted (16-7M-1997) cmy Plunkett III, Laboratory of Senetics,  Wiley (25-745) cm, 445 Menry Hall, Hadison, WI 53706, USA.  BITTLE Submission  JOURNAL  JOUR	AUTHORS	Blattner, P.R., Plunkett, G. III, Bloch, C.A., Perna, R.T., Burland, V.
Ham. S. and Sho.Y.  The complete genome sequence of Escherichia coli K-12  JUDUNIA  Selence 277 (5331), 1453-1474 (1997)  MIFFERENCE  MIFF		
TITLE JUBBLE Science 277 (5371), 1453-1474 (1997)  BYTHERMED SCIENCE SUBMISSION  DISCRETE SUBMISSION  BISTER CONSTRUCTED STREET STREET STREET STREET STREET STREET SCIENCE SUBMISSION  BUTHER CONTROL SUBMISSION  BUTHER CONTROL SUBMISSION  JOURNAL SUBMISSION  SUBMISSION SUBMIS		William Color of the Color of t
WEDLIH: PUBMED PUBMED REFERENCE 2 (residues 1 to 427) NOTIONAL NOTIONAL Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Senetics, Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Senetics, Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Senetics, Windlerty by Guy Sconsin, 445 Menry Hall, Madison, WI 53706, USA.  REFERENCE ANTHORS ANT		The complete genome segmence of Pscheriobia and in the
PUBMED RYPERENS District (residues 1 to 427) District NUMBERS District (residues 1 to 427) REPERENCE (residues 1 to 427) RIPERENCE		301EMOE 2// (3331/, 1433-1474 (1997)
REFERENCE NOTIONS NOTIONAL NOT		97426617
AUTHORS  Distance, F.R.  DIVERNAL  Submitted (16-JARN-1997) Sup Plumkett III, Laboratory of Senetics, women to Submitted (16-JARN-1997) Supplement 10-262-2534 7ax:  Submitted (16-JARN-1997) Supplement 508-262-2534 7ax:  Submitted (18-JARN-1997) Supplement 508-262-2534 7ax:  Supplement 508-262-263-2534 7ax:  Supplement 508-262-263-2534 7ax:  Supplement 508-262-263-2534 7ax:  Supplement 508-262-263-2534 7ax:  Supplement 608-262-263-263-263-2534 7ax:  Supplement 608-262-263-263-263-2534 7ax:  Supplement 608-262-263-263-263-263-263-263-263-263-263		
JUNEAU Direct Submission  Submitted (16-ARM-1997) Sup Plumkett III, Laboratory of Senetics, Wisc. 10-11-12-12-12-12-12-12-12-12-12-12-12-12-		
SUMMITTED (16-7AM-1997) cup Plumkett III, Laboratory of Senetics, University of Visconsin, 445 Menry Hall, Madison, WI 53706, USA.  REPPERENCE AUTHORS	TITLE	Direct Submission
mail: cooligenetics.wisc.du Phone: 608-262-2534 7ax:  608-265-7459  REPERENCE 3 (residues 1 to 427)  AUTHORS Blattner, F.R.  Comment of Comment of Comments of Com	JOURNAL	Submitted (16-JRR-1997) cum Plumbatt TTT
REFERENCE		
REFERENCE RATIONS Blackmer, F.R. Direct Submission Direct Submission Direct Submission Numbersity of Visconsin, 445 Menry Wall, Badison, WI 53705, USA. Email: ecolidgenetics.wisc.edu Phone: 608-262-254 Fax: COMPENIC		water evolution to the contract of the contrac
ANTHORS Blattner, F. R.  JOURNEL  Submitted (02-SEP-1997) Cup Plumkett III, Laboratory of Genetics, bushitted (02-SEP-1997) Cup Plumkett III, Laboratory of Genetics, bushitted (02-SEP-1997) Cup Plumkett III, Laboratory of Genetics, bushitted (02-SEP-1997) Cup Plumkett, G. III.  Purper Submitted (02-SEP-1997) Cup Plumkett III, Laboratory of Genetics, University of Genetics of Company of September 1998) Laboratory of Genetics, University of Plumkett, G. IIII.  JOURNEL  JOURNEL  COMMENT		
JOURNAL  Submitted (02-SEP-1997) Guy Plumkett III, Laboratory of Senetics, University of Visconsin, 445 Menny Pall, Hadison, WI 53706, USA.  Zmail: coolidgenetics.wiso.edu Phone: 608-262-2534 Fax:  608-262-7459  AUTHORS  AUTHORS  AUTHORS  AUTHORS  JOURNAL  Submitted (13-007-1998) Laboratory of Senetics, University of Visconsin, 445 Menny Ball, tadison, WI 53706, USA  This sequence was determined by the E. coli Geneme Project at the University of Visconsin-Madison (Frederick R. Blattner, director).  Supported by MIN grants MOD301 and MED1428 (from the Muman Genome Project and Committed Comm		Blattner, F.R.
Email: ecoliggenetics.wiso.edm Phone: 608-262-253 fax:  608-263-7459  A (residues 1 to 427)  MUTHORS Plunkett.s. III.  DIRPORT PLUNKET.S. III.  DIRPORT PLUNKET.S. III.  BIRGOT STATES  BIRGOT STATES  BIRGOT STATES  BIRGOT STATES  FOR A STATES  COMMENT  COMMENT  This sequence was determined by the E. coli Genet. Project at the University of Wisconsin-Madison (Frederick R. Blatter, director).  Supported by MIN grants M600301 and M601428 (from the human Genome Project and McHord). The entire sequence was independently determined from E. India KIZ strain Mc1555. Predicted open reading from the series of the entire sequence was independently determined from E. India KIZ strain Mc1555. Predicted open reading from the was a conversely described by Mark Borodovsky, Georgia Instituted States and Park Revolution of the State Open reading from the was a conversely described by Mark Borodovsky, Georgia Instituted to being amnotated with C6 site Mos., unique ID nos. for the genes in the E. coli centic Stock center (c5C) database at Pale University, Kindly supplied by Min Mark Borodovsky, Georgia Instituted loci are being amnotated with C6 site Mos., unique ID nos. for the genes in the E. coli centic Stock center (c5C) database at Pale University, Kindly supplied by Min Mark Borodovsky, Georgia Instituted to being amnotated with C6 site Mos., unique ID nos. for the database is accessible (http://gsc.biology.version of the database is accessible (http://www.genetics.wice.ed). ""  (http://www.genetics.w		DIFECT Submission
Email: ecoliggenetics.wiso.edm Phone: 608-262-253 fax:  608-263-7459  A (residues 1 to 427)  MUTHORS Plunkett.s. III.  DIRPORT PLUNKET.S. III.  DIRPORT PLUNKET.S. III.  BIRGOT STATES  BIRGOT STATES  BIRGOT STATES  BIRGOT STATES  FOR A STATES  COMMENT  COMMENT  This sequence was determined by the E. coli Genet. Project at the University of Wisconsin-Madison (Frederick R. Blatter, director).  Supported by MIN grants M600301 and M601428 (from the human Genome Project and McHord). The entire sequence was independently determined from E. India KIZ strain Mc1555. Predicted open reading from the series of the entire sequence was independently determined from E. India KIZ strain Mc1555. Predicted open reading from the was a conversely described by Mark Borodovsky, Georgia Instituted States and Park Revolution of the State Open reading from the was a conversely described by Mark Borodovsky, Georgia Instituted to being amnotated with C6 site Mos., unique ID nos. for the genes in the E. coli centic Stock center (c5C) database at Pale University, Kindly supplied by Min Mark Borodovsky, Georgia Instituted loci are being amnotated with C6 site Mos., unique ID nos. for the genes in the E. coli centic Stock center (c5C) database at Pale University, Kindly supplied by Min Mark Borodovsky, Georgia Instituted to being amnotated with C6 site Mos., unique ID nos. for the database is accessible (http://gsc.biology.version of the database is accessible (http://www.genetics.wice.ed). ""  (http://www.genetics.w	O O DECIME.	Substitute (02-529-1997) Guy Plunkett III, Laboratory of Genetics,
REFERENCE RITIORS PUNCHES RITIORS PUNCHES RITIORS PUNCHES RITIORS PUNCHES RITIORS PUNCHES RITIORS PUNCH RITIORS PUNCH RITIORS PUNCH RITIORS PUNCH RITIORS PUNCH RITIOR RITIORS PUNCH RITIOR RIT		Email: ecolifornetics wise ed Phone 500-252-254 VI 53706, USA.
ANTHORS SITIE  Direct Submission  Submitted (13-our-1998) Laboratory of Genetics, University of  Wisconsin, 445 Menry Ball, Hadison, WI 53706, USA  This Sequence was determined by the E. coli Genome Project at the  University of Wisconsin-Hadison (Frederick R. Blatter, director).  Supported by Bull grants H600301 and H601428 (from the Human Genome  Property of Wisconsin-Hadison (Frederick R. Blatter, director).  Supported by Bull grants H600301 and H601428 (from the Human Genome  Description of the Human Genome  Research of the Human Genome		608-263-7459
JOURNAL Submitted (13-ott-1998) Laboratory of Genetics, University of Visconsin, 445 Menny Mall, Madison, VI 53706, USA  COMMENT  COMENT  COMMENT		
SUBMINITE (13-oct-1938) Laboratory of cenetics, University of Visconsin, 445 Menry Mall, Madison, WI S2706, USA This sequence was determined by the E. coli Genome Project at the University of Visconsin-Hadison (Frederick R. Blattner, director). Supported by MIH grants M600301 and M601428 (from the Muman Genome Project and McMCR). The entire sequence was independently determined from E. coli KIZ strain W61555. Predicted open reading Mark Borodovsky, Seorgia generals software, kindly supplied by Mark Seorgia generals software, kindly su		
This sequence was detamined by the cooli Genome Project at the University of Wisconsin-Haddley (Frederick R. Blattner, director). Supported by MIM grants H600301 and M001428 (from the Human Genome Project and RCHGR). The entire solidate from the Human Genome Project and RCHGR). The entire solidate independently determined from E. coli K12 strain M01655. Pindependently wark Borodovsky, Georgia Institute of Technology, Rtlantage by Mark Borodovsky, Georgia Mark Boro	JOURNAL	Submitted (13-OCT-1998) Laboratory of comption and included the substitute of the su
University of wisconsin-Hadison (Frederick R. Blattner, director).  Syphother and Roth grants H800301 and H801428 (from the human Genome Project and Roth grants H800301 and H801428 (from the human Genome Prosect and Roth grants H800301 and H801428 (from the human Genome Project and Roth grants Hadison Roth grants were determined wis strain H801855. Fredicted open reading frames were determined wis strain H801855. Fredicted open reading frames where determined wis strain H801855. Fredicted open reading frames that Borodovsky, Georgia Institute H818 (and pen reading frames that have been correlated with genetic loci appen reading frames that it is stock center (cssc) database at vale Windowsky, H818 (and pen reading frames that it is stock center (cssc) database at Vale Winversity, solid senter stock center (cssc) database at Vale Winversity, solid senter is stock center (cssc) database at Vale Winversity, solid senter and congoing task whose goal is to make the genome sequence nore useful by oorrelating it with other data. Comments to the authors are appreciated. Updated information will be available at the Z. coli Genome Project's World Wide Web site (http://www.genetics.wisc.edu). *** The Z. coli K12 sequence and its amnotations are periodically updated; this is version H54. H8 sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courters of Bonnica Riley; added promoters, protein binding sites, and repeated sequence considered with a considerable of the sequence of		
Supported by MIM grants M000001 and M001428 (from the Human Genome Project and M0HUSD). The entire sequence was independently determined from E. coli entire sequence was independently determined from E. coli entire sequence was independently frames were determined using standards software. kindly supplied by wark Borodovsky, Georgia Institute of Technology, Atlanta, 6R, 30332 [c-mail: marridamber gatechen]. Open reading frames that have been correlated with genetic confirmation are being amnotated with consider Most of the dath are being amnotated with consider Most of the dath are being amnotated by the Stock Center (C6SC) database at Vale University, kindly supplied by Marp Berlyna A public version of the dathset is accessible (http://cgsc.biology.pale.edu). Amnotationare is accessible (http://cgsc.biology.pale.edu). Amnotationare sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the Z. coli Genome Project's World Wide Web site (http://www.genetics.wisc.edu). **** The L. coli X12 sequence and its annotations are periodically updated; this is version M54. Ro sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtery of Monica Riley; added promoters, protein binding sites, and repeated features described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.  WEATURES  J. 427  / organism-Escherichia coli K12*  / function="orf; Unknown" / product* orf, hypothetical protein*  1. 427  / organism-Escherichia coli K12*  / function="orf; Unknown" / ocded_bp="complement(2367272:52496532)*  / transl table=11 / notes-f427; sequence change joins ORFs yied and yied from earlier version*  1 mrsrlkdarv ppelterwae pagaslisty gfivylpqil dllhrlnsyw aegarqlvda  11 narelydm sagaslardh visi	Comment	AMAS SEGUENCE WAS DECEMBED ON THE P. AND
project and MCNGR). The entire sequence was independently determined from E. coli K12 strains were determined using census. Bit 1655 Predicted open reading frames were determined using census. Bit 1655 Predicted open reading frames were determined using census. Bit 1655 Predicted open reading frames were determined using census. Rindly supplied by Wark Borodovsky, Georgia Institute of sechnology, Rilanta, ca, 30332 [e-mail: mark@amber.gatech.edm]. Open adding frames that the control of the control of the control of the second of		
frames were determined using femedsark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 [e-mail: markGamber. gatech end.] Open reading frames that have been correlated with genetic lock are leading frames that Site Mos., unique ID nos. for the genes in the Z. coli Genetic Stock Center (CESC) database at Yake University, kindly supplied by Mary Berlyn. A public version of the database is accessible onyoing task whose goal is to make the genome sequence more useful by correlating it with other darke the genome sequence more useful appreciated. Updated informations to the authors are appreciated. Updated informations to be available at the Z. coli Genome Project's World Wide Welsite (http://www.genetics.wisc.edu) ***The Z. coli K12 sequence and its annotations are periodically updated; this is version M94. No sequence changes. Annotation updates; updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b assigned to each gene (protein- or labels. This should allow them to be searched for in Entrez as gene  PERTURES  FURATURES  FURATUR		Project and RCHGR) The entire remains 128 (Irom the Ruman Genome
Mark Borodovsky, Secogia Institute of Technology, Atlanta, GA, 30332 [s-mail: martRamber. gatech.edu]. Open reading frames that have been correlated with genetic loci are being amnotated with GG Site Mos. unique ID nos. for the genes in the E. coli tenetic Stock Center (C650) database at Fale University, kindly supplied by Marp Berlynn. A public version of the database is accessible (http://cgsc.biology.pale.edu). Amnotation of the genome is an onyoing task whose goal is to make the genome sequence more useful by oorrelating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide web site (http://www.genetics.wisc.edu). *** The E. coli KI2 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene  PEATURES  Source  1. 427  /organism="Escherichia coli K12" /strain="K12" /strain="W61655" /db xref="taxon:83333" /product="orf, hypothetical protein"		determined from B. coli Ki2 strain Mc1655 Predicted onen reading
30332 [e-mail: mark@amber gatech.edu]. Open reading frames that have been correlated with genetic loci are being amnotated with CG Site Hos., unique ID mos for the genes in the E. coli cenetic Stock Center (CGSC) database at vale University, kindly supplied by Mary Berlyn. A public version of the database is accessible (http://cgsc.biology pale.edu), amnotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated Updated information will be available at the E. coli (enome Project's World Wide etc.) site (http://www.genetics.wisde.du). *** The E. coli K12 sequence and its annotations are periodically updated; this is version H54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Honica Riley; added promoters: publicational assignments courtesy of Honica Riley; added promoters: protein binding sites, and repeated sequence described in feterence 1. The unique numeric identifiers beginning with a lowercase b' assigned to each gene (protein-or RRA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searohed for in Entrez as gene Hethod: conceptual translation.  PERTURES  FERTURES  FURTURES  FURTURES  Location/qualifiers  // Condition="Scherichia coli K12" / strain="%112" / strain="%112" / sub-strain="%11655" / db xrefs="taxon:83333" 1. 427 / function="orf; Unknown" / product="orf, hypothetical protein" / function="orf, hypothetical protein" / readed hp="complement(2367272:5249.6532)" / transl table=11 / note="f427; sequence change joins ORFs yieD and yieM from earlier version"  ORIGIN  1 mrsrlkdarv ppeltevenc pagsglistp gfivylpgil dllhrinspw aegarglvda 121 ntagriddh sacgalyada agaratuda 1221 ntagriddh sacgalyada agaratuda 1221 ntagriddh sacgalyada 1221 ntagriddh sacgalyada 1221 ntagriddh sacgalyada 1221 ntagriddh 1221 ntagriddh 1221 ntagriddh 1221 ntagriddh 1221 ntagriddh 1221 ntagriddh 1221		
Site Mos., unique ID nos. for the genes in the E. coli Genetic  Stock Center (CesC) database at Tale University, kindly supplied by  Mary Berlyn. A public version of the database is accessible  (http://cgsc.biology.yale.edu). Amnotation of the genome is an  ongoing task whose goal is to make the genome sequence more useful  by correlating it with other data. Comments to the authors are  appreciated. Updated information will be available at the E. coli  Genome Project's World Wide Web site  (http://www.genetics.wisc.edu). *** The E. coli K12 sequence and  its annotations are periodically updated; this is version M54. No  sequence changes. Ranotation updates: updated gene identifications  and products; all new functional assignments courtery of Bonica  Riley; added promoters, protein binding sites, and repeated  sequences described in reference 1. The unique numeric identifiers  beginning with a lowercase 'b' assigned to each gene (protein- or  RRA-encoding) are now designated as gene synonyms instead of  labels. This should allow them to be searched for in Entrez as gene  MERTURES  Source  1.427  /organism="Escherichia coli K12"  /strain="K12"  /strain="K12"  /function="orf; Unknown"  /products" orf, hypothetical protein=  1.427  /gene="yies"  /coded bp="complement(2367272:5249.6532)"  /transl table=11  /note="f427; sequence change joins ORFs pied and piem from  caller version"  ORIGIN  1 mrsrlkdarv ppelteewac pogsglistp gfivglpqil dllhrlnspw aegarqlvda  121 ntaartlwdm sangthrode site of the sequence of the se		
Stock center (CosC) database at Vale University, kindly supplied by Marp Berlyn. A public version of the database is accessible (http://cysc.biology.pale.edu). Amnotation of the genome is an onyoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (http://www.genetics.wisc.edu). *** The E. coli K12 sequence and its annotations are periodically updated; its version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Bonica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RMA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.  Wethod: conceptual translation.  PERTURES  Location/qualifiers  Location/qualifiers  Location/qualifiers  1. 427 / oryanism="Escherichia coli K12" / strain="M61855" / db xref="taxon: 83333"  Protein  1. 427 / function="orf; Unknown" / product="orf, hypothetical protein"  CDS  1. 427 / gene="piem" / coded bp="complement(2367272:5249. 6532)" / transl table=11 / note="f427; sequence change joins ORFs pieD and piem from earlier version"  1 mrsrlkdarv ppelteewnc yqusqllstp gfivqlpqil dllhrlnspw aegarqlvda 61 nstitsalht lflqrwrisl ivqattlngq lleeereqll sevqermatls gqlepiladm	_	have been correlated with genetic losi. Open reading frames that
Marp Berlyn. A public version of the database is accessible (http://cysc.biology.pale.edu). Amnotation of the genome is an onyoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the Z. coli Genome Project's World Wide Web site (http://www.genetics.wisc.edu). *** The Z. coli K12 sequence and its annotations are periodically updated; this is version M54. Ro sequence changes. Amnotation updates: updated gene identifications and products; all new functional assignments courtesy of Bonica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers heginning with a lowercase 'b' assigned to each gene (protein- or RRA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene Hethod: conceptual translation.  FERTURES  Location/qualifiers  Judy  Jerumism="Escherichia coli K12"  / organism="Escherichia coli K12"  / strain="N22"  / organism="Escherichia coli K12"  / strain="N22"  / function="orf; Unknown"  / product="orf, hypothetical protein"  1. 427  / gene="piem"  / coded bp="complement(2367272:5249. 6532)"  / transl table=11  / note="f427; sequence change joins ORFs pieD and piem from earlier version."  ORIGID  I mssrlkdarv ppelterwc pqusqllstp gfivqlpqil dllhrlnspw aegarqlvda 61 nstitsalht lflqrwrisl ivqattlngg lieeereqll sevgermatls gglepiladm		Site Hos., unique ID nos. for the genes in the E. goli cenetic
(http://cysc.biology.yale.edu). Amnotation of the genome is an ongoing task whose goal is to make the genome is an ongoing task whose goal is to make the genome sequence more useful by orrelating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (http://www.genetics.wisc.edu). *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. Ro sequence changes. Amnotation updates: updated gene identifications and products; all new functional assignments courtery of Bonica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RMA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.  PERTURES  Source  1427  /organism="Escherichia coli K12"  /sth_strain="MC1655"  /db_xref="taxon:83333"  Protein  1427  /function="orf; Unknown" /product="orf; Unknown"		
ongoing task whose goal is to make the genome sequence nore useful by ourrelating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (http://www.genetics.wisc.edu). *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesp of Bonica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.  Bethod: conceptual translation.  FERTURES  1. 427  /organism="Escherichia coli K12" /strain="K12" /strain="W61555" /bb_xref='taxon: 83333"  Protein  1. 427  /function="orf; Unknown" /product="orf; Unknown" /product="orf, hypothetical protein"  1. 427  /gene="yiem" /coded_bp="complement(2367272:52496532)" /transl_table=11 /note="f427; sequence change joins ORFs pieD and piem from earlier version"  1 mrsrlkdarv ppelteevac pygggllstp gfivglpqil dllhrlnspw aegarqlvda 61 nstitsalht lfigrwrlsi ivgattlngg lleeereqll sevgements gglepiladn		
appreciated. Updated information will be available at the E. coli senome Project's World Wide Web site (http://www.genetics.wisc.edu) *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesp of Bonica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.  Hethod: conceptual translation.  FERTURES  FERTURES  Location/qualifiers  1.427 /organism="Escherichia coli K12" /strain="K12" /strain="K12" /strain="W16555" /db xrei="taxon: 83333"  Protein  1.427 /gene="yiem" /coded_bp="complement(2367272:52496532)" /transl_table=11 /note="f427; sequence change joins ORFs yieD and yieM from carilier version"  ORIGIB  1 mrsrlkdarv ppelteevmc pqqsqllstp gfivqlpqil dllhrlnspw aegarqlvda 1 nstitsalht lflqqwris1 ivqattlngq lleeereqll sevgermtis gqlepiladm		
Genome Project's World Wide Web site  (http://www.genetios.wisc.edu) *** The E. coli K12 sequence and its amotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Bonica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene symonyms instead of labels. This should allow them to be searched for in Entrez as gene names.  Hethod: conceptual translation.  FERTURES  Location/qualifiers  1.427 /orgamism="Escherichia coli K12" /strain="K12" /strain="M61655" /db_xref="taxon:83333" /function="orf; Unknown" /product="orf, hypothetical protein" 1.427 /gene="piem" /coded bp="complement(2367272:5249.6532)" /transl_table=11 /note="f427; sequence change joins ORFs pieD and piem from earlier version"  I mrsrlkdarv ppelteewac pqqsqllstp gfivqlpqil dllhrlnspw aegarqlvda fi nstitsalht lfilgrwris1 ivqattingq lleeereqll sevgermtls gqlepiladm		by correlating it with other data. Comments to the authors are
(http://www.genetics.wisc.edg). *** The Z. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Bonica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene manes.  Bethod: conceptual translation.  PERTURES  Location/qualifiers  1427  /organism="Escherichia coli K12" /strain="M01655" /db_xref="taxon:83333"  Protein  1427  /function="orf; Unknown" /product="orf; Unknown" /product="orf; Unknown" /product="orf; Unknown" /coded bp="complement(2367272:52496532)" //transl_table=11 /note="f427; sequence change joins ORFs pieD and pieM from earlier version"  ORIGIN  1 mrsrlkdarv ppelteewac pqqsqlistp gfivqlpqil dllhrlnspw aegarqlvda 61 nstitsalht lfilgrwrisl ivqattingq lleeereqll sevgermtls gqlepiladm	i	appreciated. Updated information will be available at the E. coli
sequence changes. Anmotation updates: updated gene identifications and products; all new functional assignments courtesp of Honica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RMA-enooding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.  Bethod: conceptual translation.  FERTURES  Location/qualifiers  /organism="Escherichia coli K12" /organism="Escherichia coli K12" /strain="K12" /sub_strain="MG1655" /db_xref="taxon:83333"  Protein  1. 427 /function="orf; Unknown" /product="orf, hypothetical protein"  cos  1. 427 /gene="piem" /coded bp="complement(2367272:52496532)" /transl table=11 /note="f427; sequence change joins ours pieD and piem from earlier version"  origin  1 mrsrlkdarv ppelteewac yqqsqllsty gfivqlpqil dllhrlnspw aegarqlvda 61 nstitsalht lflqrwrls1 ivqattlnqq lleeercqll sevqermtls gqlepiladm		Wenome Project's World Wide Web site
sequence changes. Anmotation updates: updated gene identifications and products; all new functional assignments courtesp of Honica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RMA-enooding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.  Bethod: conceptual translation.  FERTURES  Location/qualifiers  /organism="Escherichia coli K12" /organism="Escherichia coli K12" /strain="K12" /sub_strain="MG1655" /db_xref="taxon:83333"  Protein  1. 427 /function="orf; Unknown" /product="orf, hypothetical protein"  cos  1. 427 /gene="piem" /coded bp="complement(2367272:52496532)" /transl table=11 /note="f427; sequence change joins ours pieD and piem from earlier version"  origin  1 mrsrlkdarv ppelteewac yqqsqllsty gfivqlpqil dllhrlnspw aegarqlvda 61 nstitsalht lflqrwrls1 ivqattlnqq lleeercqll sevqermtls gqlepiladm		its annotations are periodically material.
Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase "b" assigned to each gene (protein or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.  Bethod: conceptual translation.  FERTURES  Location/qualifiers  Location/qualifiers  Source  1427  /organism="Escherichia coli K12" /strain="K12" /sub_strain="Mc16555" /db_xref="taxon:83333"  Protein  1427  /function="orf; Unknown" /product="orf, hypothetical protein"  CDS  1427  /gene="yiem" /coded by="complement(2367272:52496532)" /transl_table=11 /note="f427; sequence change joins ORFs pieD and piem from earlier version"  ORIGIN  1 mrsrlkdarv ppelteewnc pqqsqllstp gfivqlpqil dllhrlnspw aegarqlvda 61 nstitsalht lflqrwrlsl ivyattlngg lleeercqll sevgermtls gqlepiladm		
sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene mames.  Method: conceptual translation.  FERTURES  Location/Qualifiers  J. 427  /organism="Escherichia coli K12"  /strain="K12"  /strain="W12"  /sub_strain="w01655"  /db_xref="taxon:83333"  Protein  1. 427  /function="orf; Unknown"  /product="orf, hypothetical protein"  CDS  1. 427  /gene="yieM"  /coded_by="complement(2367272:52496532)"  /transl_table="ill_note="fill_function"  /note="f427; sequence change joins ords pied and piem from earlier version"  1 mrsrlkdarv ppelteevac pqqsqllstp gfivqlpqil dllhrinspw aegarglvda 61 nstitsalht lfigrarisi ivqattlngg lleeereqll sevgeratis gqlepiladm		
beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.  Bethod: conceptual translation.  FERTURES Location/qualifiers  Source 1427  /organism="Escherichia coli K12"  /strain="K12"  /strain="W61655"  /db xref="taxon: 83333"  Protein 1427  /function="orf; Unknown"  /product="orf, hypothetical protein"  CDS 1427  /gene="yiem"  /coded bp="complement(2367272:52496532)"  /transl_table=11  /note="f427; sequence change joins ORFs pieD and piem from earlier version"  1 mrsrlkdarv ppelteevac pgqsqllstp gfivqlpqil dllhrlnspw aegarglvda 61 nstitsalht lflqrwrisi ivqattlngg lleeereqll sevqermtls gqlepiladm	•	semences described in reference and repeated
labels. This should allow them to be searched for in Entrez as gene names.  Method: conceptual translation.  FERTURES  Location/Qualifiers  1.427  /organism="Mc1655" /db_xref="taxon:83333"  Protein  1.427  /function="orf; Unknown" /product="orf, hypothetical protein"  CDS  1.427  /gene="yieH" /coded_by="complement(2367272:5249.6532)" /transl_table="ill_note="fertile from earlier version"  1 mrsrlkdarv ppelteewac pygsgllstp gfivglpgil dllhrinspw aegarglvda 61 nstitsalht lfigrwrisl ivgattlngg lleeeregll sevgermtls gglepiladm	·	beginning with a lowercase the watering in materic identifiers
mames.  Bethod: conceptual translation.  FERTURES  Location/qualifiers  1.427  /organism="Escherichia coli K12"  /strain="K12"  /strain="Be61555"  /db xref="taxon: 83333"  Protein  1.427  /function="orf; Unknown"  /product="orf, hypothetical protein"  CDS  1.427  /gene="yiem"  /coded bp="complement(2367272:52496532)"  /transl table=11  /note="f427; sequence change joins ORFs pieD and piem from earlier version"  ORIGIN  1 mrsrlkdarv ppelterwac pgqsqllstp gfivqlpqil dllhrlnspw aegarglvda 61 nstitsalht lflqrwrisl ivqattlngg lleeeregll sevgermtls gglepiladm	i	ANA-encoding) are now designated as gene swnonims instead of
Hethod: conceptual translation.  FERTURES  Location/qualifiers  1.427  /organism="Escherichia coli K12" /strain="K12" /sub_strain="W61655" /db_xref="taxon: 83333"  Protein  1.427  /function="corf; Unknown" /product="corf, hypothetical protein"  1.427  /gene="yiem" /coded bp="complement(2367272:5249.6532)" /transl_table="11 /note="f427; sequence change joins ORFs pied and piem from earlier version"  1 mrsrlkdarv ppelteevac pgqsqllstp gfivqlpqil dllhrinspw aegarglvda 61 nstitsalht lfigrarisi ivqattingg lleeeregll sevgeratis gglepiladm		to the substitution with the De Searched for in Particle of Come
Jocation/qualifiers  1.427 /organism="Escheriohia coli K12" /strain="K12" /sub_strain="W01655" /db_xref="taxon: 83333"  Protein  1.427 /function="orf; Unknown" /produot="orf, hypothetical protein"  1.427 /gene="yies" /coded_bp="complement(2367272:52496532)" /transl_table="ill_funct="form" /note="f427; sequence change joins ORFs pied and piem from earlier version"  1 mrsrlkdarv ppelteevac pqqsqllstp gfivqlpqil dllhrinspw aeqarqlvda 61 nstitsalht lfigrarisl ivqattlngg lleeereqll sevgeratls gqlepiladm		
1.427	FEATURES	Location/Qualifiers
/strain="N12" /sub_strain="M01655" /db_xref="taxon:8333"  Protein 1427 /function="orf; Unknown" /produot="orf, hypothetical protein" 1427 /gene="yiem" /coded bp="complement(2367272:52496532)" /transl_table=11 /note="f427; sequence change joins ORFs pied and piem from earlier version"  1 mrsrlkdarv ppelteevac pgqsqllstp gfivqlpqil dllhrinspw aegarglvda 61 nstitsalht lfigrwrisl ivgattlngg lleeeregll sevgeratls gglepiladm		1427
/strain="N12" /sub_strain="M01655" /db_xref="taxon:8333"  Protein 1427 /function="orf; Unknown" /produot="orf, hypothetical protein" 1427 /gene="yiem" /coded bp="complement(2367272:52496532)" /transl_table=11 /note="f427; sequence change joins ORFs pied and piem from earlier version"  1 mrsrlkdarv ppelteevac pgqsqllstp gfivqlpqil dllhrinspw aegarglvda 61 nstitsalht lfigrwrisl ivgattlngg lleeeregll sevgeratls gglepiladm		/organism="Escherichia coli K12"
/db_xref="taxon: 83333"  1.427 /function="orf; Unknown" /product="orf, hypothetical protein"  1.427 /gene="yieu" /coded_bp="complement(2367272:52496532)" /transl_table=11 /note="f427; sequence change joins ORFs pieD and piem from earlier version"  1 mrsrlkdarv ppelteevac pgqsgllstp gfivglpqil dllhrinspw aegarglvda 61 nstitsalht lflgrwrisl ivgattingg lieeeregll sevgeratls gglepiladm		/strain="Ki2"
Protein 1427  /function="orf; Unknown" /produot="orf, hypothetical protein"  1427  /gene="yiem" /coded_by="complement(2367272:52496532)" /transl_table=11 /note="f427; sequence change joins ORFs pieD and piem from earlier version"  ORIGIN  1 mrsrlkdarv ppelteewac pgqsgllstp gfivglpgil dllhrinspw aegarglvda 61 nstitsalht lfigrwrisi vygattlngg lleeeregll sevgermtls gglepiladm		/db xref="tayon-83332"
/function="orf; Unknown" /product="orf, hypothetical protein"  1427 /gene="pieH" /coded_bp="complement(2367272:52496532)" /transl_table=11 /note="f427; sequence change joins ORFs pieD and pieM from earlier version"  origin  1 mrsrlkdarv ppelteevac pyqsgllstp gfivglpqil dllhrinspw aegarglvda 61 nstitsalht lflgrwrisi ivgathungg lieeeregll sevgeratis gglepiladm	Protein	1427
cDs 1427 /gene="pieH" /coded_bp="complement(2367272:52496532)" /transl_table=11 /note="f427; sequence change joins ORFs pieD and pieM from earlier version"  1 mrsrlkdarv ppelteevmc pqqsqllstp qfivqlpqil dllhrlnspw aegarqlvda 61 nstitsalht lflqrwrisi ivqattingq lieeereqll sevqermtls gqlepiladm		/function="orf; Unknown"
/gene"yiem" /coded bp="complement(2367272:52496532)" /transl table=11 /note="f427; sequence change joins ORFs pied and piem from earlier version"  ORIGIN  1 mrsrlkdarv ppelteevac pgqsgllstp gfivglpqil dllhrinspw aegarglvda 61 nstitsalht lfigrwrisl ivgattlngg lleeeregll sevgeratls gglepiladm	cne	/product="orf, hypothetical protein"
/coded by="complement(2367272:52496532)"  /transl_table=11 /note="f427; sequence change joins ORFs pied and piem from earlier version"  ORIGIN  1 mrsrlkdarv ppelteevac pygsgllstp gfivglpgil dllhrinspw aegarglvda 61 nstitsalht lflgrwrisl ivgattlngg lleeeregll sevgeratls gglepiladm	CU3	1427
origin  i mrsrlkdarv ppelteevac pqqsqllstp qfivqlpqil dllhrinspw aeqarqlvda  61 nstitsalht lflqrwrisl ivqattlnqq lleeereqll sevqeratls qqlepiladn		/coded_by="complement(2367272.5249 6532)"
ORIGIN  1 mrsrlkdarv ppelteevmc pqqsqllstp qfivqlpqil dllhrlnspw aegarqlvda 61 nstitsalht lfiqrwrlsl ivqattingg lieeeregll sevgermtls gqlepiladn 121 ntaagrlwdm sagglkrgdg gliydringg lieeeregll sevgermtls gqlepiladn		/ VX =1151 CAD1E=11
ORIGIN  1 mrsrlkdarv ppelteevmc pqqsqllstp qfivqlpqil dllhrlnspw aegarqlvda 61 nstitsalht lfiqrwrlsl ivqattingg lieeeregll sevgermtls gqlepiladn 121 ntaagrlwdm sagglkrgdg gliydringg lieeeregll sevgermtls gqlepiladn		/note="f427; sequence change joins ORFs yieD and view from
1 mrsrlkdarv ppelteevmc pqqsqllstp qfivqlpqil dllhrlnspw aegarqlvda 61 nstitsalht lflqrwrisl ivqattlngg lleeeregll sevqermtls gqlepiladm 121 ntaagrlwdm sagglyrddm gllybraesi	ORIGIN	
121 ntaagriwdm saggikrodn glinkmassi sevgermtis gglepiladn	1 mrs	rikdary ppeltervac pggsglistp gfivglagil allhringer asserted
ACA REAGETWOR SAGGIRIGOD Glinknoofi namalbal.	61 nst	itsalht lilgrerist ivgattingg liceeregil severatis orieniladn
181 tfrimmena tomegodala goddilita nequencia equative sipradame	123 Rta 181 +6~	agriwdm saggikrgdy glivkygefl negoelkria egigtsreak siprndagne
		The state of the s
	当 100%	

## FIG. 14A



	a 444 2 2 3	No. 100 September 1		
ilic			(100%) (100%) (100%) (100%) (100%)	
	Confidence Certain 100% Certain 100% Certain 100% Certain 100%	Certain 100% Certain 100% Certain 100% Certain 100% Certain 100% Certain 100%	Certain 100% Certain 100% Certain 100% Certain 100% Certain 100% Certain 100%	Certain 100% Low/ 42%)
((Ō))				
	Method Local Local Local Local Local	Local Local Local Local Local	Local   Loca	Local Local Local Local Local
	Ain. score 158 157 156 149 146	146 143 143 144 144 145 144 145 145 145 145 145 145	141 141 138 138 107	102 102 102 98 98 97 96 60 60
	Target rgn. 6 - 144 5 - 130 5 - 130 6 - 144 6 - 130	- 130 - 130 - 130 - 124 - 124 - 86	1 - 96 1 - 96 1 - 96 1 - 96 0 - 100 6 - 152 10 - 134	7 - 143 11 - 149 18 - 145 8 - 145 4 - 158 8 - 146 13 - 57 10 - 202
	Targe Parage	5 - 130 (5 - 130 5 - 130 1 - 124 1 - 124		100 100 100 100 100 100 100 100 100 100
nent	2ct all	267 - 384 267 - 384 267 - 384 267 - 384 264 - 381 264 - 354	64 - 354 64 - 354 64 - 354 67 - 354 67 - 407 67 - 384	267 - 386 267 - 386 267 - 386 267 - 386 267 - 386 250 - 396 306 - 349
alignn	1000e 1267 1267 1267 1267 1267	264 264 264 264 264 264	264 264 267 267 267	
View alignment		F 51 55 68 7	14.6 14.6 11.3 11.3	11.8 10.7 11.8 11.8 18.5 14.4
e The	Organism Organism Ranovegicus Hsaplens Hsaplens Hsaplens Hsaplens	Hsaplens Hsaplens Hsaplens Hsaplens Hsaplens Hsaplens	Haspiens Haspiens Haspiens Haspiens Haspiens Haspiens	MA N/A Hsaplens Hsaplens Hsaplens Esaples Ecol
enom	H H H H H H H H H H H H H H H H H H H	HSB HSB HSB HSB	H.SB H.SB H.SB H.SB H.SB	HSB
Total selected: 0. View alignme Inpharmatica Genome Threader results:		EX EX	NESIL NESIL NESIL SER. S DO.	ETA11 ETA11 CTOR CTOR DOC: ON TOR
electe	Selected: 0 Select al Title  RYSTAL STRUCTURE OF RAT A1B1 INTEGDOMAIN FROM INTEGRIN CR3, MG2+ B0DOMAIN FROM INTEGRIN CR3, MA2+ B0RYSTAL STRUCTURE OF RAT A1B1 INTEG. MAC-1 I DOMAIN MAGNESIUM COMPLEX MAC-1 I DOMAIN CADMIUM COMPLEX	MAC-1 I DOMAIN CADMIUM COMPLEX MAC-1 I DOMAIN METAL FREE MAC-1 I DOMAIN METAL FREE CD11-1 DOMAIN WITH BOUND MIN- CD11A I-DOMAIN WITH BOUND MIN- CD1A I-D	COTTAL EDOMAIN WITH BOUND MAGNESIZED TO THE LEDOMAIN WITH BOUND MAGNESIZED THAT LEDOMOIN MAGNESIZED THAT LEDOMOIN MAGNESIZED TO TALL DOMAIN WITH BOUND MAGNESIZED WITH SOUND MAGNESIZED TO WESSOUTHORS OF THE WORK AS DOMOIN MAN YON WILLES WOR THE WONWILLES TO DOMAIN FROM INTEGRIN ALPHAZ-BETAT	IDOMAINE ROM INTEGRIN ALCHA - BETAT IDOMAIN FROM INTEGRIN ALCHA-BETAT IDOMAIN OF YON WILLEBRAND FACTOR ASDOMAIN OF YON WILLEBRAND FACTOR AT DOMAIN OF YON WILLEBRAND FACTOR HUNAN YON WILLEBRAND FACTOR NITROGEN REGULATORY BACTERIAL PRO COMPLEX: OF INVANTH-BASSWITH HUMAN.
otal se	N CR3	MACS-11 DOMAIN CADMIUM COMP MACS-11 DOMAIN METAL FREE TI-DOMAIN METAL FREE 14 I-DOMAIN WITH BOUND MI 14 I-DOMAIN WITH BOUND MI 14 I-DOMAIN WITH BOUND MI 15 I-DOMAIN WITH BOUND MI	DUND SUND FACT REVO	ALPH ALPH BRAN BRAN BRAN FACT ACTE
. 1	O. Itte OF I. F. OF I. F. OF I. AGN.	MINIM MINIM MITHOR MITHOR	TH BITH BITH BITH BITH BITH BITH BITH BI	IDOMAINE ROMINTECHNIALE IDOMAINE ROMINTECRIN ALP IDOMAIN OF VON WILLEBRAI ABIDOMAIN OF VON WILLEBRAI AT DOMAIN OF VON WILLEBRAI HUMAN YON WILLEBRAID FAC NITROGEN REGULATORY BACTE COMPLEX OF HUMANIH-BAS W
tein Total hits	CTURING CHARLES	MAIN WAIN	A CHEEN WAY	A INTERNATIONAL VOIN
otein Total	Selected: STRUCTU N FROM II N FROM II STRUCTE DOMAIN	1-100NI	DOMA DOMA DOMA ON W	HRON IN OF OFIR
olay hetical protein	SYSTAL S DOMAIN PYSTAL S MSTAL S MAC-11	MA MA 0114 19 -	CD11A I- CD11A I- CD11A I- IUMAN V CRYSTAL	I DOMAIN FROM INTI I DOMAIN FROM INTI I BOMAIN OF VON ASIDOMAIN OF VON HUMAN VON WILLEE INTROGEN REGULAT COMPLEX OF HUMAN
play				
Dis	1000 1000 1100 1100 1100 1100 1100 110	18HG:1 11DN:1 11DN:1 18HG:2 12CN:1 12CN:1	120PB 120PB 120PB 10GR 10GR	10C5.A 1.40X.B 1.403.B 1.1403.B 1.141Z.B 1.162.B
1: orf,			a 1.3 fa 1.4	
uence	A A A A A A A A A A A A A A A A A A A			
ASequence Dispay details:  AAC76768.1: off, hypot				中。中。。。
		SS		
		Seduences (%)	Bands and	
A FILT.		W 1 11 88 F. 11 60		~    ~

· 海	ğ			4-1	14	4	- 1 1 1 2	 	قديدة مراشورة علام			igen .	 	770 S	() () () ()	# 31 . 9:,73.	 6	5 5 5		Series Significant Significant					. Au.			. See	λ. ξ	Δĺ	· I	
	atic			不養	Evalue	2.00E-60	2.00E-71	2.00E-65	4.00E-38	5,00E-22	2.00E-44	3.00E-42	3.00E-42	200E-41	3.00F-41	3.00E-41	3.00E-41	3.00E-41	3.00E-41	3.00E-41	2,00E-40	8:00E-40	2.00E-38	5,005-38	5.00E - 38	100F-36	2.00E-36	2,00E-35	2.00E-35	남.		
					Blter   E	100	3 20	3 2		F	4 2.0	4 3.0	3.3	7 6	3 6	4	9.0	4 3.0	4 ·	9 2	1	4 8.	72	4 5.	4 6		4	4 2.0	4 2.0	4 1300	20000000	
3 7		を設			131	1	361	- 1	1		88	305	382	8 8	8 8	B	88	58	88	20.0	705	77	Ξ	595	988	2 2 2		307		E C		
	.⊑				Query ran: Target ran.	202 - 431	122 - 36	170 - 420	17 - 188	324 - 535	132 -	1				2-1	158	1-2	· '	2 - 15B	473-	. 6 - 177	90 ~ 31	441 - 695	415 - 588 En - ana	277 - ARA	491-747	126 - 307	118 - 307	124=		
					N ran	156 - 407	146 408	147 - 410		- 370	- 416	223 - 412	223 - 412	215 - 692	264 - 412	264 - 412	284 - 412	264 - 412	264 - 412	264 - 412	162 - 417	254 - 417	162 - 410	EZb - 291	199 - 367	158 - 957	162 - 423	239 - 412	231-412	238 - 410		#44 #64 ***
					%DIQue	28,0 156	220 146	23.0 147	21.0 241	25.0 151 - 370	247			<del></del>			16.0 264	₩	$\leftarrow$	16.0 264	٠,	17.0 254				91 150		14.0 239	-	BD 738		
, eq. (*)			73.4 73.4		%	32	22	2 2		<del></del>	Σ N	9		0.00			31:00	۳					=	2		3 6		~				01.7 ¥€
		ment	slts:		Ordanism	ΧE	aschil.	1 Jannascoll	0.00	Withermoautotrophicum	salas	lens	H.sapiens	Hisapiens	H.saniens -	H.saplens	lens	lens	sueli	H.Sapiens Psetluim	ntified	rofa	gare	liana	SSP.	Dropenlatie	llana	R.norvegicus	salus	saulens		
		Vjewalignment	ST reu	) tall	Orga	A.pemix	Mannaschil	Mjannascni	S.SD.	ermoau	Mimusculus	Hsaplens	HSB	Usanlans	H.S.P.	HSBI	- Hsaplens	H.saplens	Hsaplens	H.Saplens Peatloim	unidentified	S.scrofa	H.vulgare	A thallana	Sylvania	Disco	Amallana	Rinory	Manusculus	Hsa		
		ĬĀ.	I-BLA	Deselectall	2			-		Milh	-		-	1		1									33.00	N. A.	The Later	_				
			Reverse Maximised PSI-BLAST reusits:	والمحار					103				CINE		2 2		\$13 2.	盗	× .	MAC-11 DOMAIN CADMIUM COMPLEX				1901	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1				121 jeukocyte adheslon glycoprotein p150,85 alpha integrin Su.	Mary Control		Att Si
San C		0	aximis	Selectial		뺩	2200	ein mrein	HYPOTHETICAL 45.8 KDA PROTEIN SULDIOS	5			F 1	I-DOMAIN FROM INTEGRIN CRA MINZ+ BOUND	MAC-1 - DOMAIN MAGNESI IM COMPI EX	REE	MAC-1'I DOMAIN METAL FREE	MAC-1 I DOMAIN MAGNESIUM COMPLEY	MAC-11 DOMAIN CADMIUN COMPLEX	MAC-1 I DOMAIN CADMIUM COMPLEX				<ul> <li>magneslum-chelatase sobunit CHLD precursor</li> </ul>	2				lipha int			
		otal selected: 0.	rse M	Sele		452aa long hypothetical protein	*** HYPOTHETICAL PROTEIN MJ0077.	conserved hypothetical protein	301	magnesium chelatase subunit Chil			leukocyte integrin alpha chaln	2 2	N (S	MAC-1   DOMAIN METAL FREI	ETALF	MOIS	S S	S E			돌	E 5	opmit The state of the state of	Conserved professional	מומאב א	듣	50,85	erkolntearin alaha d chala	Part of the last	
341		otal se	Reve		岸	ypothet	L PROT	pomeu	KOA	elatase	Not given	Not given	egrin a	2000	MAGN	MINIM	MINIM	MAGN	CAD	CAD	Not given	COLID	Vig-protoporphyrin I>	e sobla	Mg chelatase subunit	Compelved profess	27G7.20	alpha D Integrin	oteln p	adala n		マイ を た
4 / 28.				508.		long h	ETICA	refraels	AL 45.6	slum ch	ž	ž	5 C		MAN		r'i DON	MAIN	OMAII	OMAIN	Ž		/g-prot	helatas	Ag che	ממנומפו	212	旧	glycopi	matura.		
	_			Total hits: Selected:	i i	452ae	NPOT	COUSE		magne			leuko			MAC	MAC	1 100	10			100		slum-c	2		D 88		heslon	a	Same Contra	PDB
	proteir	18 G	18. 18.	Tot			表		H P P O	1		3	100		MAC		製造	MAC	MA	ME				magne			P		cyte ad	,		<b>△</b>
	play thetical					-				-	-			1		-			3	7	.1.				_ 3	1		1	1 leuko		A TOTAL OF	S-PROT PPDB
	Disp hypoth			of 1;	Code	80016.	Q60384	AAG79810.	Q55874	184857.	CAA30479,	(58544.)	AABZ4BZ1,	200	1919	11DN:2	11DN1	1BHO:2	-1BH0:2	18HQ:1	CABGB268		CAB58178.	AAD52031.	wit	10000	AAF22885.1		AAF23482.	338547		တ [
	Ce [				-		O C	AA	Ø	AAE	Š	A	¥ľ	1	-					=  0	2	AA	CAE	. AAL	BAA167	AABBBC	AAF228	¥	AA	AABBE	TANK TANK	Š
ype	uen : 5768.1			Page	r Detail			ء اہ	10	^	۵	٦		1	10	_	۵	4	4	1		_		4	4	1	ماد	^	٥	4		▶GenBank ▶SWI
arch 1	led Sequence Display Query details: AAC76768.1: orf, hypothetical protein			。 以	Cluster	اوا	۰	۰	•	۰	۰	딕	9	4	+	1				ות	•	0	٠	ò		٥	•	•	٥	9		Ő
Edit Search Type Help			5.					÷		-		•	Sec	1		77	•	r. ,	-		Š.	- 10 € 10 € 10 €	ं कर जो	<del>† .</del>						. :		
File Ec	Aligned Query	Tools	T.	Tr.	•	•]	0		U		>	<	Seguences	c	❷		1		<b> +</b>		7. A.		Jands	289	E	9 Phe	20	· ·	<u>.</u>	,		
							_							_		_		-					_	<u> </u>		_	-		_			

SUBSTITUTE SHEET (RULE 26)

TLGEQIFSIEGGT

FIG. 15

AlEye output (December 13, 2000 3:07 PM)

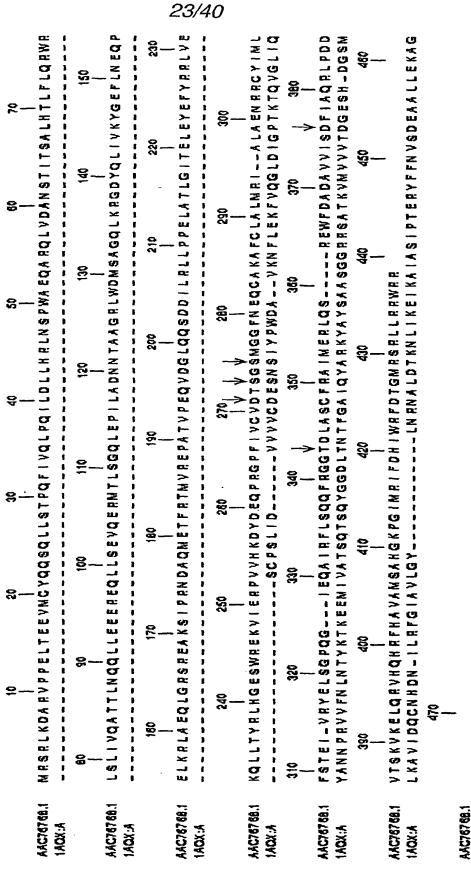
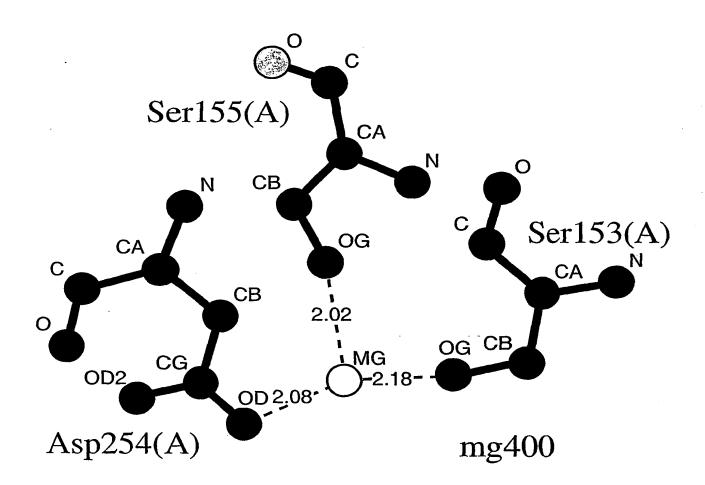
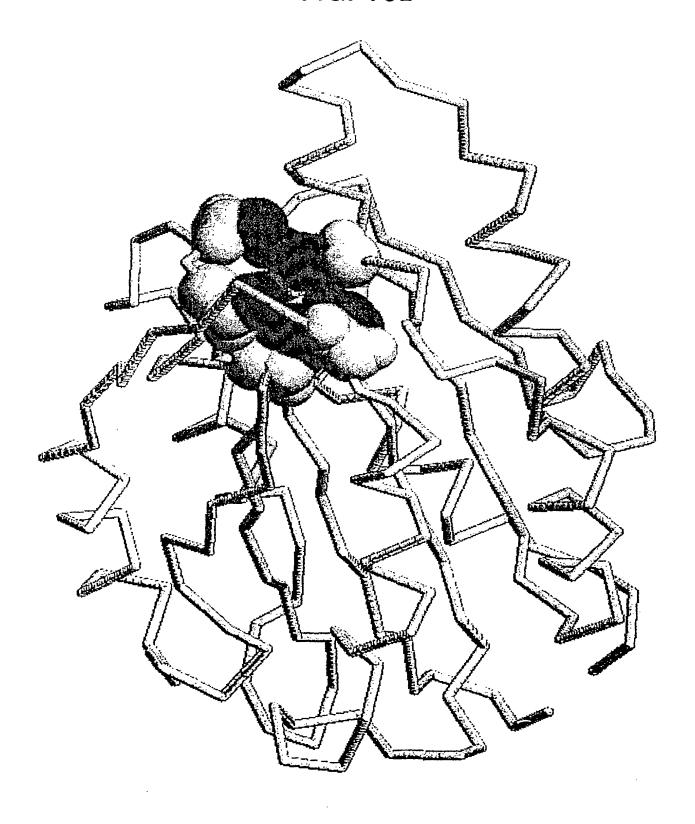


FIG. 16A



1aox: MG400

FIG. 16B



File Edit View Go Communica			生活类型技术展示企业出
Bookmarks & Location	http://victoria.inpharmati	ca.co.uk/~volker/BPD3targe	t.html
	Target Min Interface	ing inph	armatica
Soloot Vour Oue	ry Soguenoo		
Select Your Que	ary Sequence		
	number (e.g. 1 QMA): [j] J	M and chain (e.g. B):	]
OR			[ <del></del>
• Enter one Swiss-Pro	t accession (e.g. P27504)	or GenBank proteinID (e.g.	CAB08761.1):
Select Database	<b>)</b>		
Release: DEVF9 =	BPD3□		
Apply Filters			
• Iteration Filter: PSI-I	BLAST matches to be exch	ıded:	
	None		
<u> </u>			
(using the first_PB_iter and detected after 4 or more using Genome-Threade	annotation). This allows you to PSI-BLAST iterations. Matcoor rare not effected by this option and by Genome-Threader it w	o focus on more remote homo hes detected using PSI-BLAS on. However if one match is:	T with negative iterations or
□Homo sapiens	☐Rattus norvecicus (Rat)	- □Mus musculus (Mouse)	□Danio rerio (Zebra fish)
a =0=[100%	La La Contra de la Contra del Contra de la Contra del la Contra de la		

# FIG. 18A 27/40

			ocation:   http://London-brid				in/volker/	getTargetB	PD3.p	ł				
81 additional hits identified by both, Genome Threader and PSI-BLAST:  mbined Genome Threader and PSI-Blast output: PSI-BLAST values are shown in maroon!														
12list	BPD links	VA4AY link	Title	Organism	Div.	%ID (GT,PSI)	Query rgn. (GT,PSI)	Target rgn. (GT,PSI)	Aln. score (GT)	Conf. (GT)	la (te			
	AAA59544.1 drill through Top50Blas1Hits Fied.Seq.View	AAA59544.1	Not given	Homo sapiens	PRI	100%, 100% unmaskedSW	1-187, 1-187	148-334, 148-334	488	100% urmaskedGT	1			
<b>&gt;</b>	AAB24821.1 drill through Top50BlastHits Red.Seg.Vew	A A B24821.1	leukocyte integrin alpha, chain	Homo sapiens	PRI	100%, 100% unmaskedSW	1-187, 1-187	148-334, 148-334	488	100% unmaskedGT	1			
	Q99715 drill through Top50BlastHits	Q99715	COLLAGEN ALPHA 1(XII) CHAIN PRECURSOR.	Homo sapiens (Human).	PRI	28.9%, 28% unmaskedSW	2-186, 2-179	439-617, 2322-2494	456	100% unmasked GT	1			
	AAB38702.1 drill through Top50BlastHits Red.Seg View	A A B38702.1	cartilage matrix protein	Homo sapiens	PRI	28.9%, 25% unmaskedSW		274-452, 40-221	446	100% urmasked GT	1			
ـــــ	AACD1506.1 drill through Top50BlastHits Red.Seg.View	AAC01506.1	type XII collagen	Homo sapiens	PRI	28.4%, 28% unmaskedSW		137-318, 137-318	445	100% unmasked GT	1			
J	CAA72402.1 drill through Top50BlastHits Red, Seg, View	CA A72402.1	collagen type XIV	Homo sapiens	PRI	28.7%, 30% unmaskedSW		6-185, 6-185	442	100% unmaskedGT	1			
 ز	AAB38547.1 drill through Top50BlastHits Red, Seg, View	<u>A A B38547.1</u>	leukolntegrin alpha d chain	Homo sapiens	PRI	61%,60% unmaskedSW	1-187, 1-187	148-334, 148-334	439	100% unmaskedGT	1			
	CAB71222.1 drill through Top50Blas1Hits	CAB71222.1	dJ238D15.1 (collagen, type Xff, alpha 1)	Homo sapiens	PRI	27.1%, 22% unmaskedSW	1-186, 2-186	293-472, 1430-1620	439	100% unmaskedGT	1			
	CAA07569.1 dril through Top50BlastHits	CAA07569.1	matriin—4	omoH ensique	PRi	27.9%, 24% unmasked SW	2-186, 3-187	344-525, 34-215	418	100% unmaskedGT	2			
J	CAB46380.1 drill through Top50BlastHits	CAB46380.1	dJ453C12.3 (matrān-4)	Homo sapiens	PRI	27.9%, 24% unmaskedSW	2-186, 3-186	385-566, 34-214	418	100% ammaskedGT	2			

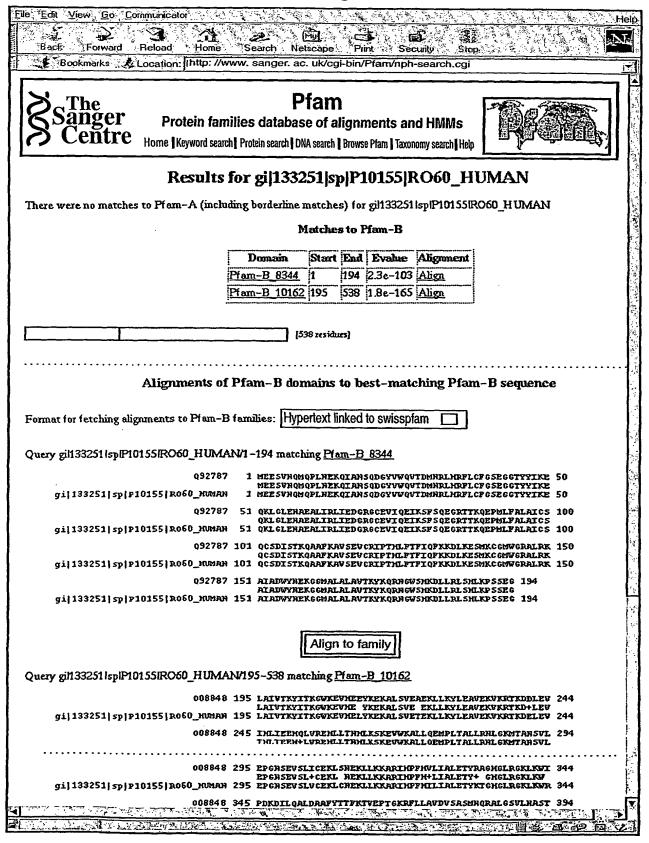
윤	$\mathbf{Z}$	1		8:3			isid:				1271	Ťi	2 <u>k</u>	y gan	MAY 1				KEE.			
			unmaskedGI reverse Hit	89 60% urmasked01	89.85% unmaskedGT	87,73% unmaskedGI	85 59% unmaskedGI	B4.45% urmaskedGI	BO.79% unmaskedGI reverse Hit	Z4.2% urmaskedGT	72.84% unmaskedGI	ZI 48% UnmaskedGI	Z0.11.2 unmaskedG1	68.75% unmaskedGI reverse Hit	67 41% vnmpskedGI	67.41% urmaskedGI	SB.06% urmaskedGT	63.4% uomasked <u>GI</u>	63.4% unmaakedGI	unmaskedGI reverse Hit	59,48%, yrmaskedGT	S6.92% Unmasked@I
2			99	18 18	В	78	78	4	29	88	69	28	88	22	63	<u>.</u>	99	99	88	٤	63	R
			390-503	20-105	405-474	147-289	318-403	134-275	384-497	442-552	363-429	373-504	373-504	749-851	578-647	682-801	475-643	475-643	123-308	750-852	152-204	130-166
			73-185	88-9	73~143	2-115	89-9	7-141	73-185	84~185	76-143	6-145	6-145	22-122	124-188	6-9t	30-185 4	30-165	20-184	22-122	2-55	2-36
			WS bestemny	MSBakeroun	26.8% urmaskedSW	13.1% urmasked 3W	MS payeeus	VZ 24.	203. vomsvkedSW	18.8% wmasked 8W	WSDaylemun 25%	11.1% unmarked SW	U. 1.Z. urmaskedSW	18.5x vronsskedSW	20.8% urmaskedSW	从Spayeemun	15.4% unmaskedSW	15.4% VormaskedSW	MSpayseum	18.5%	16.42 unmasked3W	24.3% unmaskedSW
			PRI	Œ.	Œ	Ē	E	Æ	E	Œ	Ē	Ē	Æ	Œ	Ē	Œ.	Œ	Œ	EE &	Ĕ	РВІ	E.
			Homo saplens	Homo sapiens	Homo saplens	Homo saplens (Human).	Homo saplens	Homo saplens	Homo saplens	Homo saplens	Homo sapiens	Homo saplens	Homo saplens (Human).	Homo sapiens	Homo sapiens	Homo saplens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapkns
	, and o	bridge.inpharmatica.co.uk:/cgi-bir	KIAA0639 protein	G7c protetn	N-copine	INTEGRIN BETA-8 PRECURSOR.	L NG37	Sec23 protein	T copine !	. polymerae	discorte ((shake to KIAAGSSE))	Not given	60 KDA RO PROTEIN (80 KDA RIBONUCLEOPROTEIN RO) (RORNP) (SJOGREN SYNDROME TYPE A ANTIGEN (85-A)).	L texokinas !	KIAA1439 protein	Sec248 protein	protocacherin gamma A11 anort form proteth	Drotocather'n gamma A11	T (senexin XXXI)	hexokhase I.	hisgrin beta-7 subunit	hitoph varant betalf
	Figure 3.	ittp://london	8AA31611.1	CAB52192.1	BAA75899.1	P26012	AAD21820.1	CA465775,1	AAC15920.1	AAC63290.1	CA887610,1	AAA35532.1	PIOLSS	AAC15863.1	BA492872.1	CAA10335.1	AA043756.1	AAD43714.1	CA40833.1	AAA52646.1	AAAS9185.1	AAB65421.1.
Communicator	Relead	. Location: Th	drill through Igp50BlastHis	CABS2192.1 dril fivough Tap50BastHis	8AA75939.1 drill Intough IQD508las11413.	P25012 drill Incount Top508ts1Hits	AAD21820.1 dril through Top50BastHis	CA465775.1 dril incoxen Top508tes IFfts	AAC15920.1 drii through Top50BlastHits	AAC63290.1 ora through Top50BastHis	CA887810.1 drif inrough Top50BastHits	AAA3SS32.1 drii ibrounh IopSOBaathis	P10155 drift (two uph Top50 Blast Hits	AAC 15883.1 drill through Top50Biss I His	9AA92872.1 drill incough Top50Basil His	CAA10335.1 drill inrough Top50BasiHite	AAD43766.1 dril ihrough Tor508haiHita	AAD43714.1 drill iprovab Top50Bas His	CAA08933.1 dril through Top50BleatHis	AAAS2848.1 drin itrough Top50Biss Hris	dril Incount full Incount Ion50BastHis	drii treuch
ברוו דופא מס	Back Forward		Red.Seg.Vew	Red Seg Mew	Red Seg Mew	Red Sen Ver	Red. Seg. Mew	Red.Seg.New	Red Sea Mew	Red. Seg. Mew.	Red.Sen.Vew	Red Seg Mew	Red.Seg.Vew	Red Seg Mew	Red Sea New	Bed. Sea, Vew	Red.Sea.Maw.	Bed. Seg. Mew	Red.Sea.Vew	Red, Seg, Mew	Ard Sea Vew	Red. Sea. View
5 91	8	77	7	٦	7	٦.	٦	7	ר	7	7	٦	λ	7	٦.	'n	7	7	7	7	7	
_		1::::	<u></u>										===	<u> </u>		<u> </u>	<u></u>		<u> </u>	<u> </u>	<u></u>	

FIG. 18E

29/40

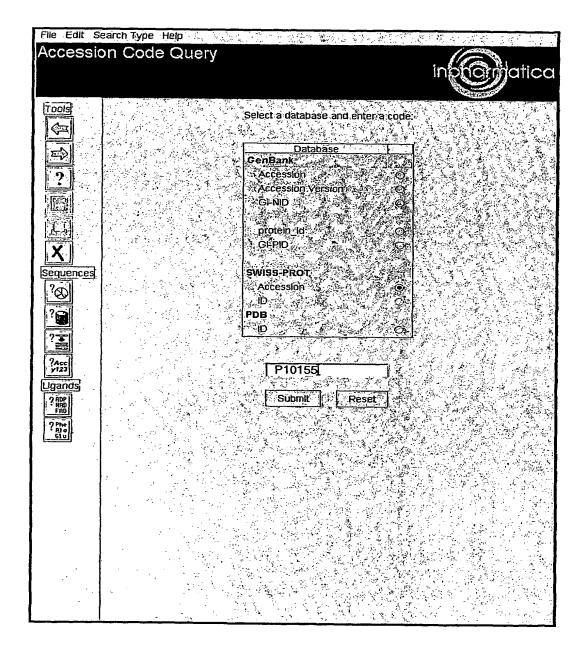
	19		Massassassass		i di saddi manggi		L
- )		1		7777			ā. ā. 🕄
				100 - 100 -	100 100 100 100 100 100 100 100 100 100	혦	
				42.24.24.25.25.25.25.25.25.25.25.25.25.25.25.25.		S 1912	
Tarana a			201121 - 121		17. 17. 17. 17. 17. 17. 17. 17. 17. 17.	1 - CAAS	
			17 17 17 17 17 17 17 17 17 17 17 17 17 1	77.75.75.75.75.75.75.75.75.75.75.75.75.7	AND THE PROPERTY OF THE PROPER	A95324	
			100 - 100 -	150 150 150 150 150 150 150 150 150 150		v- Tam	
			2004111 20111111111111111111111111111111	1. SAKE 1. SAKE 1. SAKE 1. SAE	100 100 100 100 100 100 100 100 100 100	7. T	TO VENT
Ģ.			A A STATE OF THE S	CK42.7 ACDIGO ACO ACDIG	100 100 100 100 100 100 100 100 100 100	this thisper triciopy tillispp thisper teneral thisper trisper.	AANSTER AANSTE
			4 - 100 H	4. 110 4.		7. 1535	10077777 177777777777777777777777777777
			100 100 100 100 100 100 100 100 100 100	17-17-17-17-17-17-17-17-17-17-17-17-17-1	100 - 100 -	W'Ta	
			A. CAAS AAFISA A	A-440 A-440	17-61-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-	PANTA PANTA	AARINA AA
			17.77.77.77.77.77.77.77.77.77.77.77.77.7	70. 71. 70. 71. 71. 71. 71. 71. 71. 71. 71. 71. 71.		THEST	1100
25,000			20011111111111111111111111111111111111	100 100 100 100 100 100 100 100 100 100	100 100 100 100 100 100 100 100 100 100	यःभा	2333333 333333
			AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	A LUMBER OF THE PROPERTY OF TH	100 AU 10	SAA2	AALTHUR THE
			100 100 100 100 100 100 100 100 100 100	17 17 17 17 17 17 17 17 17 17 17 17 17 1	THE PARTY OF THE P	7173777	
	Stone Stone		1000 - 20	7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7		75	
	ं व	PD3.pl	1000 - 10	7. LOAN 1. CANDA 1. C	11.042.1 12.042.1 12.042.1 12.042.1 12.042.1 12.042.1 12.042.1 12.042.1 13.	77.4827	OAAL OAAL OAAL OAAL OAAL OAAL OAAL OAAL
	© gods	lang lang	11. 2011 1. 20	14.04.0 14.04.0 10.04.0 10.04.0 10.04.0 10.04.0 10.04.0 10.04.0	PANIST OF THE PARIST OF THE PA	TACTOM!	AATAI AAHAH AAHAH AAHAH AAHAH AAHAH
6	নুহা ও	et l'ar	AA211316 AA211316 AA211316 AA211316 AA211316 AA21316 AA21316 AA31316	11.15.45.15.15.15.15.15.15.15.15.15.15.15.15.15	100 - 100 -	P-T1811	7 TH
	Security	Jker/g	19 19 19 19 19 19 19 19 19 19 19 19 19 1			77	
	S	Vuid	10.000	1.1022 1.102 1.102 1.1022 1.10	AAANI AAANI	**************************************	
	m E	k/g	TOTAL THE COLUMN TO THE COLUMN	100 - 100 M	AND THE SECOND STREET OF THE S	TSIRRE	
	, u a	a.00.1	100 100 100 100 100 100 100 100 100 100	2021-12-12-12-12-12-12-12-12-12-12-12-12-1			
	Netscape	maje Dig	ALCANIA ACTION OF THE PROPERTY	Cretion Partition		rations': .AAruuu	44444
		and I		itive it	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	adve ite	
3	gich (	-bridge.inpharmatica.co.uk./cgi-bin/volker/getTargetBl		sing 'postave iteratons'.  Sing bostave lasted cast the lasted la		ing 'negadve iterations'. Vilter - adrivita, dapirii	
	Š	ggu	100 100 100 100 100 100 100 100 100 100	fied usi		ed usir L. Gazz	
7	<b>₹</b> 5.5	http://london-	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	identi	AAANINA AAANIN	identifi	24414 14414
ē	***	P H		S Vere		A WES	
Communicator	Relogd (C.)	Bookmarks & Location		match Tribut Tri		23 out of these 649 PSI-BLAST matches were identified us easurus, gaaltect, garruns, gaarress, garrens, ga	A2) Genome Threader Matches: 2023 muches found by Genome Threading Alexand Annual Annu
Comp	É	L.		I.AST		.AST.	Water Parkers of the control of the
යි	Forward	इ				18-18d	A2) Genome Threader Matches: 2925 muches found by Genome Thread ASSESSED AND ASSESSED
View View	h .	okma	= = = = = = = = = = = = = = = = = = =	३ व्यवस्थानस्य ३ व्यवस्थानस्य	######################################	10.4.00 10.4.00	4 B 20111111
富	きる意	المد		CALLER CA		of thei	more and the second sec
哥哥		Bookmarks . Location		626 out of these 649 PSI -BLAST matches were identified to the control of the con		23 out	A2) G 2025 mu
뜨	/	F	· <del></del>		·		750

File Edit, View Go. Communicator Help
See Back Fawward-, Religied Home Search Netscape Print Security Shop Stop.
: Bookmaiks 🏌 Location: //london-bidge/inphaimatica/BPDDEVT/cgi-bin/tsp.pl/?ktep_esn=55817&password=carss_app&usemama=carss_app&usede_sid=BPDDEVT 💌 個一一個一個
Aligned annotation view for P10155 (downloading image)
1 50 100 150 200 250 300 350 400 450 500 Primary database information:
Z Prosite match
Toron marion marion
Source databank Accession code P10155 Gene name SSA2 OR RO60 Download sequence in FASTA format
Links to other resources: Wew custom hyperlinks
Search in ExPASy  European Bioinformatics Institute: IIPR000504  Online Mendelian Inheritence in Man(234700*)  Online Mendelian Inheritence in Man(50063)  Protein Information Resource  A 30596
TION RO SMALL RBONUGLEOPROTEINS CONSIST OF FOUR SMALL RNA MOLECULES OF 85-
OF THE ADDROVED THE STATE OF TH



File Edit View	Go Communicator	alc.
	ward Reload Home Search Netscape Print Security Stop	
Bookmail	ks 💹 Location: http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?cmd=Retrieve&db=Prot	1
LOCUS DEFINITION ACCESSION	ROGO_HUMAH 538 aa PRI 01-FEB-1996 60 KD RO PROTEIN (60 KD RIBOHUCLEOPROTEIN RO) (RORNP) (SJOGREN 5YNDROME TYPE A ANTIGEN (55-A)). P10155	4
PID VERSION	g133251 P10155 GI:133251	i,
DBSOURCE	swissprot: locus R060_HUMAH, accession <u>P10155;</u> class: standard. created: Mar 1, 1989.	
	sequence updated: Mar 1, 1989. annotation updated: Feb 1, 1996.	H
	xrefs: gi: gi: <u>177782</u> , gi: gi: <u>177783</u> , gi: gi: <u>387656</u> , gi: gi: <u>387657</u> , gi: gi: <u>107626</u> xrefs (non-sequence databases): MIM <u>600063</u> , MIM 234700, PROSITE	
KEYWORDS	P500030 Ribonucleoprotein; RNA-binding; Systemic lupus exythematosus;	1
SOURCE ORGANISM	Antigen. human. Homo sapiens	1.00
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REPERENCE AUTHORS TITLE	1 (residues 1 to 538) Deutscher, S.L., Harley, J.B. and Keene, J.D. molecular analysis of the 60-kDa human Ro ribonucleoprotein	1
JOVRHAL MEDLINE	Proc. Natl. Acad. Sci. V.S.R. 85 (24), 9479-9483 (1988) 89071722	
REMARK REFERENCE AUTHORS	SEQUENCE FROM N.A 2 (residues 1 to 538) Ben-Chetrit,E., Gandy,B.J., Tan,E.M. and Sullivan,K.F.	4
TITLE	Isolation and characterization of a cDNR clone encoding the 60-kD component of the human SS-A/Ro ribonucleoprotein autoantigen	1
JOURHAL MEDLIHE REMARK	J. Clin. Invest. 83 (4), 1284-1292 (1989) 89198084 EQUENCE FROM N.A.	ů.
соммент	This SWISS-PROT entry is copyright. It is produced through a	
	collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. The original entry is available from http://www.expasy.ch/sprot	
	and http://www.ebi.ac.uk/sprot	7.
-	[FUNCTION] UNKNOWN. [SUBUNIT] RO SMALL RIBONUCLEOPROTEINS CONSIST OF FOUR SMALL RNA MOLECULES OF 85-112 NT, EACH OF WHICH IS COMPLEXED WITH A 60 KD	
	PROTEIN. RO RNPS MAY ALSO CONTAIN AN ADDITIONAL 52 KD PROTEIN. [SUBCELLULAR LOCATION] CYTOPLASMIC. [DISEASE] SERA FROM PATIENTS WITH SYSTEMIC LUPUS ERYTHEMATOSUS	10
	OFTEN CONTAIN ANTIBODIES THAT REACT WITH THE HORMAL CELLULAR RO PROTEIN AS IF THESE ANTIGEN WAS FOREIGN.	
FERTURES	[SIMILARITY] CONTAINS 1 RNA RECOGNITION MOTIF (RNP). [SIMILARITY] STRONG, TO XENOPUS 60 KD RO PROTEIN. Location/qualifiers	1
source	1538 /organism="Homo sapiens"	· .ÿ
Proteir	/db_xref="taxon: 9606" 1.:538 1.:538	
Region	/product="60 KD RO PROTEIN" 9398	
Region	/region_name="Domain" /note="RNA-BINDING (RHP2) (BY SIMILARITY)." 124131	
_	/region_name="Domain" /note="RNA-BIHDING (RNP1) (BY SIMILARITY)."	
Region	239 /region_name="conflict" /note="K -> R (IH REF. 2)."	A.
Region	515538 /region_name="conflict"	7
ORIGIN	/note="GMLDMCGFDTGALDVIRHFTLDMI -> ALQHTLLHKSF (IR REF. 2)."	-
67 13	esvnamap lnekqiansa dayvwqvtdm nrlhrflofg seggtyyike aklglenaea rliedgrg ceviqeiksf sqegrttkae pmlfalaios qosdistkaa afkavsevor thlftfiq fkkdlkesmk ogmwgralrk aiadwynekg gmalalavtk glarngwshk	
181 d.	Lirishikp sseglaivtk pitkgwkevh elykekalsv eteklikyle avekvkrtkd evihliee hrivrehlit nhikskevwk aliqempita lirnigkmta nsvlepgnse	:
361 fx	tveptgkr fllavdvsas mngrvlgsil nastvaaamc mvvtrtekds yvvafsdemv	<u>.                                    </u>
421 pc	pyttdmtl qqylmamsqi paggtdoslp miwaqktntp advfivftdn etfaggyhpa lreyrkkm dipaklivog mtsngftiad pddrgmldmo gfdtgaldvi rnftldmi	1.3
	。 · · · · · · · · · · · · · · · · · · ·	Y
THE STATE CONTRACTOR OF THE BUILDING STATES TO	A TABLE TO A STATE OF THE PARTY	ك

## FIG. 22A



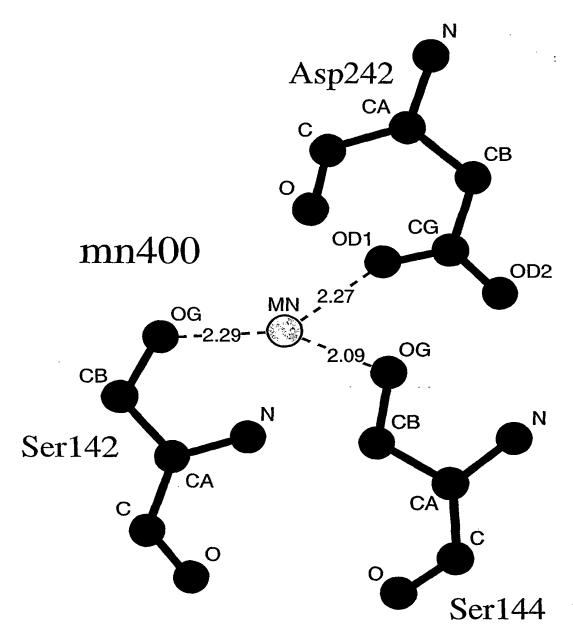
CCO		<b>जा</b> री	
Jath			
		Confidence Low(67%) L	LOW (10%)
			Cocal   Coca
(SS-A)).		Aln. score 55 55 55 55 55 55 56 56 56 66 67 67 67 67 67 67 67 67 67 67 67 67	
AANTIGEN		Emetron   120 - 250   145   6 - 145   6 - 145   6 - 145   6 - 145   6 - 145   6 - 145   6 - 145   6 - 145   6 - 145   6 - 145   6 - 145   6 - 145   6 - 145   6 - 145   6 - 145   6 - 145   102 - 248   6 - 143   103 - 252   120 - 254	可以上數上層為當品樣品當品樣上性極度變無關
)   3 y (60 KDA RIBONUCLEOPROTEIN RO) (RORNP) (SJORGEN SYNDROME TYPE A ANTIGEN (SS-A)).  Total selected: 1.	Inpharmatica Genome Threader results.  Select all Deselect all.	42 172 873 - 504	504 504 505 504 507 172 130 130 130 171 171 171
EN SYNDR	ne Threader resu Deselect all	2610 10.9 11.7 11.7 10.3 10.3 10.3 11.5 10.9 10.9	103 103 105 105 105 105 105 105 105 105 105 105
лр) (sJORG	smatica Genor Select all	निवास किया है।	Asapiens Fulsease Fulsease Fulsease Hsaplens Tichuzi Syphimurium Syphimurium Syphimurium Siphimurium Siphimurium Siphimurium Siphimurium Siphimurium Militari
EIN RO) (ROFINP) Total selected:	Inpharmal	TING  CRYO-CRYSTALLOGRAPHY OF A.TRUE.SU.  1-DOMAIN FROM INTEGRIN CRS. MN24 BO.  MAC-1 I DOMAIN MAGNESIUM COMPLEX  MAC-1 I DOMAIN MAGNESIUM COMPLEX  MAC-1 I DOMAIN CADMILM COMPLEX  MAC-1 I DOMAIN CADMILM COMPLEX  MAC-1 I DOMAIN CADMILM COMPLEX  MAC-1 I DOMAIN METAL FREE  TRP SYNTHASE (B60N-IPP-SER) WITH K+  1-DOMAIN FROM INTEGRIN CRS. MG2+ BO.  CRYSTAL STRUCTURES OF MUTANT (BETA.  ITRYPTOPHAN SYNTHASE (ECA2.120) FRO.  TRYPTOPHAN SYNTHASE (ECA2.120) FRO.  TRYPTOPHAN SYNTHASE (ECA2.120) FRO.  TRYPTOPHAN SYNTHASE (ECA2.120) FRO.  TRYPTOPHAN SYNTHASE (ECA2.120) FRO.	LONSTAL STALE OF THE YOW WILLEN TO STALE TO STALE THE COMPLEX OF AN FABLETING TO STALE THE COMPLEX OF AN FABLE OF THE COMPLEX OF AN FABLE TRUCTURE OF THE CALLED IN TARPOPHAN SYNTHASE (E.C.42.1.20) IN TARPOPHAN STALE STRUCTURE OF YEAST KARNOPHAN STALE STRUCTURE OF THE MOESIN FECONSTAL STRUCT
OPROTEIN Tot	.s: 31	CRYO-CRYSTALLOGRAPHY.OF A TRUE SUI I-DOMAIN FROM INTEGRIN CR3, MN2+ BO. MAC-1 I DOMAIN MAGNESIUM COMPLEX MAC-1 I DOMAIN MAGNESIUM COMPLEX MAC-1 I DOMAIN METAL FREE MAC-1 I DOMAIN CADMIUM COMPLEX MAC-1 I DOMAIN METAL FREE TRP SYNTHASE (D60N-IPP-SER) WITH K+ -DOMAIN FROM INTEGRIN CR3, MG2+ BO. RYSTAL STRUCTURES OF MITANT (BETA. IDOMAIN FROM INTEGRIN ALPHAZ-BETA IRYPTOPHAN SYNTHASE (ECA2.120) FROMERYPT OPHAN SYNTHASE (ECA2.120) FROM INTEGRIN ALPHAZ-BETA	CHAISTAL STANDER OF THE VOW WILLE.  TOWN AND WINDS THE CONNELEX OF AN FAB.  FOOTAND MOUTH DISEASE WINDS TYPE.  FOOTAND MOUTH DISEASE WINDS TYPE.  AT DOMAIN OF VON WILL EBRAND FACTOR CRYSTAL STRUCTURE OF THROSINE AMINI. TRYPTOPHAN SYNTHASE (E.C.42.120) IN T. TRYPTOPHAN SYNTHASE (E.C.42.120) IN T. TRYPTOPHAN SYNTHASE (E.C.42.120) IN T. ASPARTATE AMINOTRANSFERASE (FC.26. CRYSTAL STRUCTURE OF YEAST KARYOP.  SPARTATE AMINOTRANSFERASE, YZZSER. ASPARTATE AMINOTRANSFERASE, YZZSER. CRYSTAL STRUCTURE OF THE MOESIN FE. CRYSTAL STRUCTURE OF THE MOESIN FE. CRYSTAL STRUCTURE OF THE MOESIN FE.
IBONUCLE	Total hits Selected:	TITILE CRYO-CRYSTALLOGRAPH DOMAIN FROM INTEGRAPH MAC-11 DOMAIN MAGIN MAC-11 DOMAIN MAGIN MAC-11 DOMAIN MAGIN MAC-11 DOMAIN MAGIN MAC-11 DOMAIN CADN TRP SYNTHASE (DEGN-IP TRP SYNTHASE (DEGN-IP TRYSTAL STRUCTURES OF TRYSTAL STRUCTURES OF TRYSTAL STRUCTURES OF TRYSTOFHAN SYNTHASE TRYSTOFHAN SYNTHASE TRYSTOFHAN SYNTHASE TRYSTOFHAN SYNTHASE TRYSTOFHAN SYNTHASE	ASTAN CHISTAN INCOLORS  AGORI I DOMAIN FROM INTE  IGGCI STRUCTURE OF THE CHEND AT DOMAIN OF VON V  ITON AT DOMAIN OF VON V  ITON AT RYPTOPHAN SYNTHA  ITON AT RYPTOPHAN SYNTHA  ITON A RAPARTATE AMINOTE  IEERA CRYSTAL STRUCTURE  IEERA CRYSTAL STRUCTURE  IARHE ASPARTATE AMINOTE  IARHE ASPARTATE AMINOTE  IEFTE CRYSTAL STRUCTURE  IEFTE CRYSTAL STRUC
	3	CRYO-CRYSTALL I-DOMAIN FROM MAC-1 I DOMA MAC	DOWAINE STRUCTURE TO DOMAIN THYPTOPH SPARTAT SPARTAT SPARTAT SPARTAT SPARTAT SPARTAT SPARTAT
Ce Dis	1 of	145A:A 115A:A 110A:2 110A:2 110A:1 110A:1 110A:1 110A:3 110B:3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	AOXIVA AOXIVA 1 AUG 1 1 AUG 1 AUG
KOA KOA	Page		
ale Seaton ype ned Sequel Query details: P10155: 60 KD/			
Aligne Que	个 别		2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2

	U				11	1	. <u> </u>	· · · · · · · · · · · · · · · · · · ·	<u>.</u>			_2 }. 		a			Tana	eriori est. Vilorita		23. 23	V Triv	7	जिल
	i i i		宝山		41, 57, 41,	***					ありま		¥	1	· · · · · · · · · · · · · · · · · · ·							* * * * * * * * * * * * * * * * * * *	
	TEN.					ن حا			- - I-	# J		160	r.±		1		1				THE STATE OF THE S		
					E value	0,00E+00	0.005+00	0.005	0,000-+00 0,000-+00		0.00E+00	1.00E-126	6.00E-25	6.00E-05	2.00E-04	7.00E-04	n.nF-na						
						5	5	3/8	5 C		ō	=	6	19	202	12	15						
	-				n. 18.18	_			1	-   ~	2	-	6	60	1	3	1_						
'''					arget rgn	- 516	-23/		200	54 = 649	15 - 531	64 - 639	273 - 372	513-627	517 - 627	1-429	507 - 569			\. \. \. \. \. \. \. \. \. \. \. \. \. \			4
}	SS-A))					(0)		-	-	╁	╁	-	╌				7 50						
	GEN (				Query rgn.	1-516	- 23 -	- 538	100	- 538	- 5 88	1 - 537	266 - 381	264 - 379	264 - 375	34 - 32	264 - 327						
	A ANT	t t	::	数数 数据引	0  Q %	83.0	2 2 2 3	0.08		36.0	340	17.0	26.0   2	21.0.12	23.0 23	27.0.72	25.0 26					37	
	play N (60 KDA RIBONUCLEOPROTEIN RO) (RORNP) (SJORGEN SYNDROME TYPE A ANTIGEN (SS-A))	View alignment	Reverse Maximised PSI-BLAST results:	ĠĪ.		Ť		2	1.	+-	<del>i</del>	,	<del>-</del>	<del>!</del>	i								
	30ME	new a	AST	Desellect all	Organism	Hsaplens	Manusculus	Vimusculus Vitaliile	Hsaolens	C.elegans	Dradiodurans	<b>Dmelanogaster</b>	Scoelicolor	Mimusculus	<b>Anorvegicus</b>	Thermophila	Hsaplens			W		1. 1000 1000 1000 1000 1000 1000 1000 1	
	SYNDF		S-BI	Jesel	Ö	Ĭ	<b>S</b> :	N A	1	S	Drac	Ome	Sa		Rino	Time	7.5						
	GEN (	1	sed P													16.00 10.00				4		i sa Mez	
	Jors)	-	aximi mix	Select all													ļ.,			74.00			
1	ORNP)	Total selected: 1.	se M	Selec											nent 1	8	등						<b>新</b>
	10) (R	sele	Jever 1			<u></u>	<u>.</u>			5	E	ς.	hypothetical protein	telomerase protein-1	elomerase protein component	telomerase component p80	elomerase-associated protein T						
	TEINE	Tota		13	<b>a</b>	Not given	Not given	rivor given	Not given	Not given	Not given	Not given	etical	ase pr	rotein	com	sociate						
	ОРЯО							- la			2	2	ab de la constant de	elomer	rase p	merası	se-as		A.				
	IUCLE			Total hits: Selected:							į			7	telome	telo	omera		1			構造	SS-PROT PPDB
	RIBON		.	Sel		j.			.]					ў. Э.			<u></u>						
	ау							1															PHO
				of 1.		532.1	3 5	je		241.1	833.1	876.1	98.1	89	590.1	- 69 -	107.1						ISSI
	Dis			ŤIII	9 3	AAA3553	AACESTA	AAC3B00	AAB8155	CAA9824	AAF 10833.1	AAF 4587	CAB4640	AAC53043.	AAB5169	AAC4660	AAC5110						S
Hei	<b>NCE</b> AROF			Page 1	Ħ					0						٦	<u>√</u>						▼ ►GenBank ►SWIS
Type	due 18: 60 KD			E III	Juster Detail	+	1	1			1	-+	-	1				ر ا ا					enBe
Search Type Help	l Sequence Dis ry detalis: P1015s. 60 KDA RO PROTEI	<u> </u>		Ö		1	u 	0	°	•	٥	0	<u> </u>	<u> </u>	<u>^</u> ا	^	•				4		
	igned Sequence Query details: P10155. 60 KDA RO PI			· <del></del>							F	8		:			7	4	<u>سين</u>	- <del> </del>	17		
e Edit	ign	Tools F	1,	î.	<b>5</b>	7		E	A	<u></u>	7	Sednences	Ø		<u>.</u>		<b> -</b>	新	gands	急量品	2 Phe		₫ <sub>V</sub>
윤	N A		!		_	<u>.</u> #,	===	<u> </u>		_	<u> </u>			<u> </u>		111		1115				!! 	

FIG. 23 AlEye output (January 2, 2001 1:29 PM

8	, <u>8</u> —.a	1 g-0	6/40 ! ൠ̃—≖	1 400 - 400 - 400	- <del>8</del> !	- 280 - 280	. 12.
VIQE-K®	ADWYNEK	i F	ト		カ > ! !!	PADVF!V AFK!LVV <sub>S</sub>	RN FT LOM I QNQ LREKI
5-8 00 10	150 150 184 L R K A J	230         	<u> </u>	9 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	. 64 	18 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	64 LDV 1 RB - A L K T 1 QB
50   	1 <u>5</u> — ₹	220 220 3 V E T E K L	300 	380 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	L PM   WAG V R E L FN   540 	OMCG FOT QVNN FEL
KIGLEN	7 7 7 7 7 7 8		SVLEPG	 		A G G T D C S A T G L A K V	₩ 2 2 2 2 2 3 4 4 4 4 4 4 4 4 4 4 4 4 4 4
50 GTYY - KEQ	130 H.F.T.F.19	210 	LGKMT	,	! <del>Y-&gt;</del> .	2	ELN TIAGE
40 	120 	200     V T K Y   T	<b></b> .	360 7	40	8 L V K P   T 520 	
30 	110 		270   WKALLQEM	350     LKALDAA	00-	FONN PN P	LI VCGMTSIVIGVGDAF
1 7 DW 7 Y D O	2000 ST K	1	KS KEY	LKWBPDEE			KMD I PAKL
20 	180 180 – 18	180 180 180 180 180 180 180 180 180 180	260 	340 140 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	450	# 1 E K D & 4 C	1. 人にお日子 吊ドンション・ファート ロートロート
≥ Q (	90 100   100	ALAL	250   H L ! E E H	330   	÷-	A A MCM V V T R T E K D M Y V V A M E Q L K K S K T L F S L N Q Y S E E F R - H 450	TÖN ET FAGG YH PAIAL BEYRKM Tog ek fgo Plg y ed vipeladre
P1015S M						S MIL	P101SS T (

FIG. 24A



1jlm: MN 400

FIG. 24B

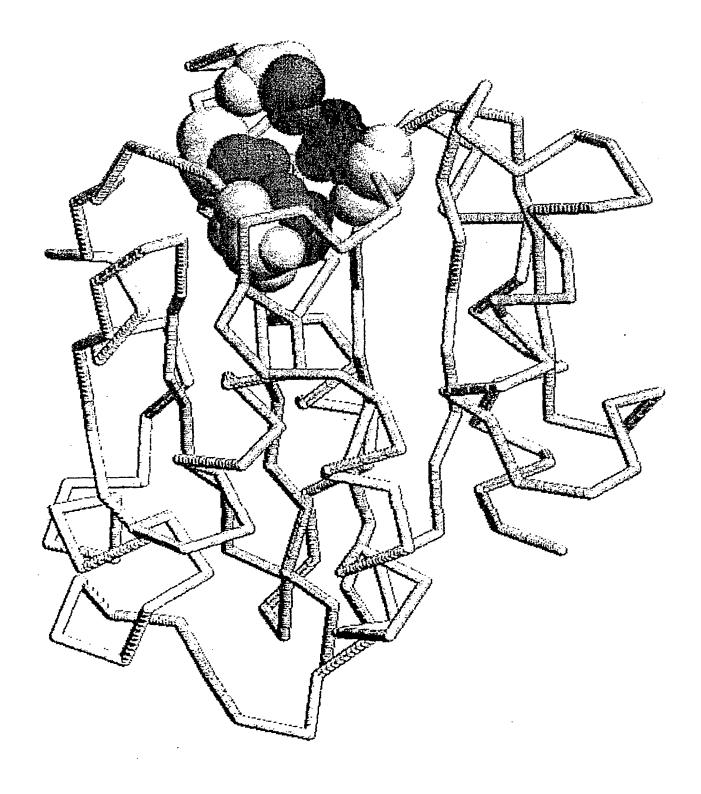


FIG. 25
AlEye output (January 4, 2001 3:18 PM)

	ć	39/40			
70 	150 MLFALA ( C S Q C S D 'L L F A L A V C S Q C S D 'T L F A L A V C S Q C S D M I F T L A I C A R I S T	230 IGRALRKAIADWYN IGRALRKAVADWYN IGRALRKAVSDWYN	310 	380 VWKALLQ-EMPLT VWKALLQ-EMPLT IWKSLLQ-DMPLT VWETLFDVSMPMT	460 FDEELLKALDAAF PDKDILQALDAAF PDTSIVEALDNAF PNOKICDALEAGF
60 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	140 SOEGRITKEP SOEGRIAKOEP SOEGRAAKOEP NAENRNPKMNA	220 KCGMW KCGMW KCGMW KCGMW	290 	370 HILLTNHLKOKE HILLTNHLKOKE HILLTIHLKOKE	450   GHGLRGKLKWR GHGLRGKLKWI GHGNRGKLRWI GRGDKGSLTWE
50 	130 GCEVIOEIKSF GCEVIOEIKSF GCEVVOEIKTF GHMVLKELALI	210 1 OFKKDLKESMI 1 OFKKDLKESMI 1 OFKKDLKEGMI 1 OFKKDLKEGMI	280 	360 .   1   1 E E H R L V R E   H L I D E Y R L V R E   A A I K K H G L V R E	440   1   1   1   A   E T Y K T   H V C   A   E T Y R A   H   L V A   E T Y K K   N   L T A R A V Y A Q
H R M M M M M M M M M M M M M M M M M M	120 Alirliedor Alirliedor Allriiedor Riikiiegor	90 CRIPTHLFTF CRIPTHLFTF CRIPTHLFTF CRIPTHLFTF	270 LKPSSEGLAI LKPSSEGLAI LKPANEGLTM MDSRSHGGSE	350 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	430   
30 00 1 1 1 0 00 1 1 1 0 00 1 1 1 1 0 00 1 1 1 1	EOKLGLENAE EOKLGLENAE EKKLGOENAE EKKLGOENAE	80 - AAFKAVSEV - AAFKAVPEV - AAFRAVPEV DSALDLIPEV	260 DLLRLSH DLLRLSH DLLRLSH	340 	420   L V C E K L C N E K I   L I C E K L S N E K I   S V C E R L T N E E K I
20 00 00 00 00 00 00 00 00 00 00 00 00 0	08 EGGTYY!K GSEGGTYY!K GSEGGTYY!K GSEGGTYY!K	170 	N N N N H H H H H H H H H	330 	410 1 S V L E P G N S E V S L 1 S V L E P G N S E V S L 5 S V L A P A S S E V S L 7 G A L D E K R V D I
MEESVN MEGSVN	V TOMN BLHAFLOF V TOMN - V RAFLOF V SO DETO V RAFLOF	160   STKQ	250 	320 Y K E K A L S V	400 
P10155 AAF19049.1 AAG38001.1 CAA98241.1	P10155 AAF19049.1 AAC38001.1 CAA98241.1	P10155 AAF19049.1 AAG38001.1 CAA98241.1	P10155 AAF19049,1 AAC38001.1 CAA89241.1	P10155 AAF19049.1 AAC38001.1 CAA88241.1	P10155 AAF19049.1 AAC38001.1 CAA88241.1

FIG. 25 (contd.)

AlEye output (January 4, 2001 3:18 PM)

